AD-A034 730 HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA F/G 5/10
READING AND READABILITY RESEARCH IN THE ARMED SERVICES.(U)
SEP 76 T G STICHT, D W ZAPF N00014-76-C-0312
UNCLASSIFIED HUMRRO-FR-WD(CA)-76-4 NL

1 OF 4
AD-A
034 730

U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A034 730

READING AND READABILITY RESEARCH IN THE ARMED SERVICES

HUMAN RESOURCES RESEARCH ORGANIZATION
ALEXANDRIA, VIRGINIA

SEPTEMBER 1976

026068

Final Report

HumRRO
FR-WD-CA-76-4

ADA034730

HumRRO

Reading and Readability Research in the Armed Services

Thomas G. Sticht
and
Diana Welty Zapf
Editors

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street • Alexandria, Virginia 22314

September 1976

Prepared for:

- Air Force Office of Scientific Research
- Army Research Institute for the Behavioral and Social Sciences
- Navy Personnel Research and Development Center
- Office of Naval Research

Approved for public release, distribution unlimited.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: HumRRO FR-WD-CA-76-4	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Reading and Readability Research in the Armed Services	5. TYPE OF REPORT & PERIOD COVERED: Final Report
	6. PERFORMING ORG. REPORT NUMBER: FR-WD-CA-76-4
7. AUTHOR(s): Thomas G. Sticht Diana Welty Zapf	8. CONTRACT OR GRANT NUMBER(s): N00014-76-C-0312
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 154-385
11. CONTROLLING OFFICE NAME AND ADDRESS: Personnel and Training Research Office of Naval Research (Code N00014) 800 N. Quincy Street, Arlington VA 22217	12. REPORT DATE: September 1976
	13. NUMBER OF PAGES: 316
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Jointly sponsored by Air Force Office of Scientific Research, Army Research Institute for the Behavioral and Social Sciences, Navy Personnel Research and Development Center, and Office of Naval Research.
Research performed by HumRRO Western Division (CA), under Work Unit READWRITE.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Comprehensibility	Literacy	Text design
Education	Programmed instruction	Useability
Functional literacy	Readability	Writing
Information processing	Reading	
Job reading	Technical documentation	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report contains the proceedings of the Conference on Reading and Readability Research in the Armed Services, held on October 28-30, 1975. The Conference brought together reading and technical writing/readability experts from civilian R&D centers with R&D specialists and operational personnel within the Armed Services for focused discussion on reading and readability/text design problems in the Services. This report presents the perspective papers prepared by research personnel representing the Air Force, Army, and Navy; the comments and suggestions by the civilian consultants; and recommendations

20. *(Continued)*

for future R&D in the Services. The major recommendations stressed the following needs:

(1) Need for each Service to develop a comprehensive plan leading toward a total career development program.

(2) Need for functional, job-related literacy training.

(3) Need to consider oracy skills as well as literacy skills.

(4) Need for more basic research in the areas of reading, readability, comprehensibility, and useability.

SUMMARY AND RECOMMENDATIONS

MILITARY PROBLEM

Whenever the job or training performance of military personnel is adversely affected by inadequate literacy skills, the Armed Services face a literacy problem. The Services cannot rely upon the nonacceptance or limited assignments of marginal literates to avoid this problem. Times of increased mobilization or lower economic opportunity will bring a certain percentage of marginal literates into the Services, as will considerations of equal opportunity and upward mobility. Furthermore, the increasingly technical nature of military documentation demands high literacy skill levels; even personnel of moderate reading ability may fall short of these high levels.

The overall "literacy problem," which involves both oral and written language skills (oracy and literacy), has two sides: (1) the *personnel* side - problems of assessment of the literacy and oracy skills of personnel and of attempts to raise those skills through literacy training; and (2) the *materials* side - problems of assessment of the reading difficulty (readability, comprehensibility, useability) of job and training materials and of attempts to lower the difficulty of those materials and to otherwise modify job and training requirements to reduce their literacy demands.

RESEARCH AND DEVELOPMENT PROBLEM

The Armed Services have active research and development efforts which address both sides of the literacy problem. Considerable R&D activity on this problem also exists outside the military. Increased interaction between Armed Services and civilian R&D personnel would provide opportunity for fresh insight and ideas for future R&D activities within the Services. To meet such an objective, an agreement was made to hold a Conference which would bring together key military and civilian personnel in these R&D areas.

APPROACH

A *Conference on Reading and Readability Research in the Armed Services* provided an opportunity for an exchange of ideas between military R&D personnel and civilian consultants. In this interchange, past and present Armed Services R&D efforts, delivered in perspective papers, were discussed in Working Group sessions which met to consider recommendations for future R&D. Historical papers were delivered by J. Dexter Fletcher (NPRDC), Jack Hiller (ARI), and James R. Burkett (AFHRL). Papers on current research on the personnel side of the literacy problem were delivered by Thomas Duffy (NPRDC), John Caylor (HumRRO), and Steven Groff (AFHRL). Papers on the materials side of the literacy problem were prepared by Thomas Curran (NPRDC), Arthur Siegel (Applied Psychological Services), Richard Kern (ARI), Robert Johnson (AFHRL), and William Muller (NAVAIRSYSCOM).

The civilian consultants, who led the Working Group sessions, were Jeanne Chall (Harvard), Samuel Gibbon (Children's Television Workshop), Robert Glaser (LRDC), John Guthrie (IRA), George Klare (Ohio University), Michael Macdonald-Ross (Open University), Ernst Rothkopf (Bell Labs), and Richard Venezky (University of Wisconsin). Macdonald-Ross presented the Conference keynote address, "Research and the Transformer," which gives a unique perspective to R&D efforts. This volume presents a record of those Conference activities.

MAJOR RECOMMENDATIONS

While the Conference generated many specific recommendations and suggestions for future R&D, these recommendations may be grouped into four main concerns.

(1) The need for each Service to develop a comprehensive plan leading toward a total career development program. Such a program would incorporate R&D efforts on both the personnel and the material sides of the literacy problem and would provide the individual with access to literacy training and job training, when needed, throughout his career development.

(2) The need for functional, job-related literacy training. Unlike general literacy training, functional literacy training can integrate into job training and on-the-job experiences, providing motivation in terms of successful job-related performance.

(3) The need to consider oracy skills as well as literacy skills. Oral language skills can affect the nature and amount of literacy training, and also job training, that will be most effective for an individual. Much research on oracy skills remains to be done.

(4) The need for more basic research in the areas of reading, readability, comprehensibility, and useability. Such research would increase our knowledge of the processes involved in reading and reading training; the increased understanding would, in turn, help us to produce both better readers and better reading materials.

FOREWORD

In October 1975 a Conference on Reading and Readability Research in the Armed Services was held at the U.S. Naval Postgraduate School in Monterey, California. The Conference brought together representatives from research and development organizations within the U.S. Air Force, Army, and Navy, operational personnel from these services, and civilian scientists from major universities and other research and development organizations to discuss mutual concerns on issues of reading assessment and training, and the design of more readable and useable technical documentation. This report presents the proceedings of the Conference and additional papers resulting from post-Conference deliberations.

The Conference was funded by all three armed services. Specific organizations, and people in those organizations whose activities made the Conference possible are:

Air Force Office of Scientific Research, Dr. Alfred R. Fregly

Air Force Human Resources Laboratory, Dr. James R. Burkett

Army Research Institute
for the Behavioral and Social Sciences, Dr. J. E. Uhlaner

Navy Personnel Research and Development Center, Dr. Edwin Aiken

Office of Naval Research, Dr. Marshall J. Farr

In addition to the foregoing persons, Dr. Joseph Ward of the Army Research Institute for the Behavioral and Social Sciences contributed considerably to the development of interest in and support for the Conference.

The task of serving as technical monitor for the project fell to Dr. Marshall J. Farr of the Office of Naval Research. In this role he served as the primary contact for HumRRO personnel, coordinated contracting and Conference planning meetings among the various service organizations, and provided counsel and assistance on various matters pertaining to the conduct and reporting of the Conference.

Rear Admiral Isham Linder, Superintendent, U.S. Naval Postgraduate School, kindly made the facilities of that institution available for the Conference. Captain Dean Taylor, Jr., Chief of Staff of the USNPGS, acted as host and provided support for the Conference. Mrs. Ruth Guthrie, Conference Coordinator at the USNPGS, performed valuable services in the arrangement of facilities and services for the Conference.

The Conference was conducted by the Human Resources Research Organization, Western Division, where Dr. Howard H. McFann is Director. Dr. Thomas G. Sticht was Principal Investigator for the project. Mrs. Diana Welty Zapf served as Conference Coordinator and Associate Editor of the Conference proceedings.

The work was performed under Department of the Navy Contract N00014-76-C-0312 issued by the Office of Naval Research under Contract Authority NR 154-385, with funding support supplied by the Office of Naval Research, the Navy Personnel Research and Development Center, the Army Research Institute for the Behavioral and Social Sciences, and the Air Force Office of Scientific Research.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

PART I

INTRODUCTION

THE CONFERENCE ON READING & READABILITY RESEARCH IN THE ARMED SERVICES

BACKGROUND TO THE CONFERENCE

The Literacy Problem: Producing Better Readers

In the midst of Project 100,000 in November of 1970, a *Working Group on Listening and Reading in the Armed Services* convened at the Presidio of Monterey, California. The purpose of this Working Group was to study the literacy problems created for the Armed Services by the large influx of less literate men of Project 100,000.

As a consequence of its deliberations, the Working Group prepared a recommendation for literacy research and development in the Armed Services and submitted this recommendation to Dr. Ralph Canter, then Director of Research, Office of the Assistant Secretary of Defense, Manpower and Reserve Affairs.

In April of 1974, the Manpower Development Division of the Air Force Human Resources Laboratory released a *Report On Literacy Training Programs in the Armed Services* (McGoff & Harding, 1973). This report contains a Prologue signed by the then Deputy Assistant Secretary of Defense for Education, Mr. M. Richard Rose, who comments on the problem of literacy as it continues beyond Project 100,000 and into the All Volunteer Force. This Prologue states:

> It appears highly probable that the estimates of a continued flow of accessions with reading problems into the Armed Forces are essentially correct . . .
>
> The need for a more comprehensive effort on this problem is evident in the unanimous recommendations of the Working Group on Listening and Reading in the Armed Services made in November 1970. This recommendation was that 'literacy training be designed following a system approach, which would include the thorough assessment of literacy requirements of the various military occupations, the orderly structuring of training programs geared to satisfying the occupational requirements, and, *most importantly*, well designed evaluative procedures to provide feedback for program improvement' . . .
>
> Progress in dealing with the literacy problem is essential if the productive potential and more effective utilization of a portion of our manpower resources are to be achieved.
> (McGoff & Harding, 1973, pp 9-10.)

As this quotation indicates, as of April 1974, there was continued concern within the Armed Services with training problems posed by the need to utilize personnel having low literacy skills. Since then, the economy has stimulated many more academically capable youth to seek employment in

Preceding page blank

the Armed Services, and the problem of extremely large numbers of marginally literate personnel has at least receded, although it has not disappeared completely. Furthermore, it seems likely that if economic conditions in the nation were to improve considerably, the Services would once again face the problem of providing literacy training to accommodate less apt personnel.

Facing the foregoing problem directly, all three Services — Army, Navy, Air Force — today are vigorously researching methods for providing reading training in an optimal manner for career development of all personnel. This means that, while attention is being given to the very lowest level of reader who is likely to enter the Service, there is also concern with the development of reading training for personnel reading anywhere below the high school level. Many of the higher quality recruits entering the Services today under the press of economic insecurity are reading well below the high school level for which they are certified by diploma. Hence, even with the higher quality input into the Services today, there is a continued need for research and development on methods for providing literacy training for career development. And, indeed, such research is underway in all three Services.

The Literacy Problem: Producing Better Reading Materials

A major problem facing many military personnel is that the technical manuals, training literature, job performance aids, and other written materials that they are given are frequently of such poor quality (high difficulty, poor format, incomplete, poorly organized) as to render them practically unuseable,[1] regardless of a person's reading ability. Again, this problem has been well recognized by the three Services, and active research and development programs are currently underway to improve technical documentation. The success of these R&D efforts should lead to reduced reading demands of technical materials in many military jobs, and hence, render such jobs more accessible to less literate personnel (as well as producing materials more useable by all personnel).

The R&D Problem

Thus, the Armed Services have research and development programs which take two main approaches to literacy problems: R&D to *produce better readers* through literacy training, and R&D to *produce better reading materials*, i.e., to reduce the difficulty levels of manuals, texts, and other printed materials. However, in recent conversations with scientists in each of the Services, it was acknowledged that there is currently a great deal of research and development activity outside the Armed Services on problems of literacy and design of textual materials which has relevance to the problems found in the Armed Services. A need was expressed by researchers within the Department of Defense to communicate with experts on

[1] In this report, we use the spelling *useability* rather than *usability*. We believe that the former spelling helps preserve the semantic relation between *use* and *useability*, a relation of major importance in improving materials.

literacy and readability from civilian institutions. It was felt that such discussions might provide fresh, creative insights for long-term solutions to the recalcitrant problems of literacy and improvement of technical writing which the Services face.

A Conference on Reading and Readability Research in the Armed Services

Because of the expressed desire by many Armed Service R&D personnel to meet with and interact with civilian literacy experts, as well as to interact with each other in response to conversations with civilian experts, the Human Resources Research Organization (HumRRO) proposed to conduct a *Conference on Reading and Readability Research in the Armed Services*. Such a Conference would bring together Armed Services R&D personnel, concerned military operational personnel, and civilian experts in literacy and technical writing to focus specifically upon literacy and technical writing R&D problems within the Armed Services.

This proposal was discussed with representatives of the Office of Naval Research, the Navy Personnel Research and Development Center, the Air Force Human Resources Laboratory/Technical Training Division, and the Army Research Institute for the Behavioral and Social Sciences. During these meetings, basic agreement on the need for such a Conference was reached, and it was also suggested that, following the Conference, a working group composed of representatives from the R&D organizations of the Navy, Army, and Air Force and the Conference Chairman from HumRRO should meet to consider the proceedings of the Conference and to study recommendations for R&D in reading and readability from the Conference for integration into the Conference Proceedings for wide-scale dissemination.

As a consequence of the negotiation meetings with the four R&D organizations from the Armed Services, the Office of Naval Research, Personnel and Training Research Programs Division, was authorized to act on behalf of the three Services and to contract with HumRRO to conduct the *Conference on Reading and Readability Research in the Armed Services*.

OBJECTIVES AND FORMAT OF THE CONFERENCE

Major Objectives of the Conference

The goal of the Conference, that of exchanging ideas in consideration of the course of reading and readability R&D, is embodied in these major objectives:

1. To bring together military R&D and operational personnel with civilian experts in the fields of reading and readability of technical writing to discuss the Services' problems in these areas; to discuss what the Services have done, what they are currently doing, and what each Service is anticipating for immediate next steps in these areas; and to exchange ideas and insights for additional actions which might be taken to pursue solutions to the problems of literacy and technical writing encountered by the Services.

2. To develop, through small-group discussions of key issues and idea papers produced by the civilian experts based upon their reflections of the Conference, information which would be reviewed in post-Conference, tri-Service meetings to develop recommendations for future R&D in reading and readability.

3. To produce a published document containing the proceedings of the Conference, including the recommendations evaluated in the tri-Service, post-Conference meetings, which would bring the Armed Services' concerns with literacy and readability of technical writing to the attention of a wide range of Department of Defense personnel and civilians interested in these vital issues.

Format of the Conference

The *Conference on Reading and Readability Research in the Armed Services* was held on October 28-30, 1975, at the Naval Postgraduate School in Monterey, California. The mornings were devoted to presentations of perspective papers prepared by research personnel for and from the Air Force, Army, and Navy. The papers dealt with (1) *past* R&D on literacy training and readability, (2) *ongoing* research on literacy training, (3) *ongoing* research on improving the readability of technical writing, and (4) *ongoing* research on improving the useability of technical writing through consideration of factors other than those involved in readability research (e.g., format, organization, selection of content via task analysis). Following each paper, a non-Service related civilian expert in reading or technical documentation (readability, design of texts) commented on the paper. These experts served as consultants to the Conference and were available to Service representatives throughout the Conference. The perspective papers and the comments by the civilian consultants are contained in Part II of these Proceedings.

In the afternoon sessions, the expert consultants to the Conference conducted five Working Groups dealing with these R&D topics: (1) readability, (2) design of texts and graphic material, (3) training systems and materials, (4) literacy training programs, and (5) basic skills in reading (phonics, comprehension). These Working Group sessions discussed various issues in these topic areas and suggested recommendations for future research.

At a general meeting of conferees on the last afternoon of the Conference, The Working Group leaders (the civilian consultants) presented oral summaries of the key issues and recommendations to emerge in their groups. At this same general meeting, operational personnel from each of the three Services delivered oral statements of research and development needs from an operational perspective. The combined observations and suggestions that emerged from the Conference are presented in Part III of these Proceedings.

In addition to the papers in Parts II and III, two papers were prepared which present perspectives on literacy training and design of materials that cut across all three Services. These papers follow in the remaining pages of Part I.

LITERACY TRAINING IN THE ARMED SERVICES

Thomas G. Sticht
Human Resources Research Organization

Throughout their histories, the Armed Services have contended with the problems of training young adults who are characterized by their slowness to learn, low performance on various aptitude and classification tests, and poor performance on tests of academic skills (reading, arithmetic). Several chapters in the present volume discuss the past and current efforts within the Air Force, Army, and Navy to deal with personnel having low academic skills through the provision of remedial literacy training. Generally, the upshot of these reviews is that current literacy training efforts are inadequate for the job. The papers by Caylor (Army), Duffy (Navy), and Groff (Air Force) all point to the lack of convincing evidence for the effectiveness of short term (6-8 weeks) remedial reading programs of the type favored by the military since World War II, and which are currently in operation in the three Services.

The attempt to cope with problems resulting from the accession of personnel of low literacy skills by providing remedial literacy training is only one of four strategies which have been used by the Armed Services for dealing with this problem. The three remaining strategies include not accepting marginally literate personnel; accepting these personnel, but assigning them to jobs requiring only low levels of academic skills; and redesigning of training and job materials to accommodate the less literate personnel. As with remedial literacy training, none of these strategies has proven entirely satisfactory. The following pages provide a more detailed discussion of all four strategies and suggest reasons for the lack of satisfaction with these approaches as they have been implemented up to now.

Following the discussion of the four strategies for dealing with literacy problems, a general plan is discussed for the development of a training system which would not only provide training in technical job skills and knowledge, but which would also incorporate the teaching of basic and advanced literacy skills within the context of job skills training.

In presenting the four strategies for contending with literacy problems and the general plan for an integrated job skills/literacy skills training system, problems are discussed of a conceptual and methodological nature which will have to be resolved if any course of action based on the assessment or training of literacy skills is to be meaningfully accomplished.

STRATEGIES FOR DEALING WITH LITERACY PROBLEMS

Four strategies have been followed at various times to cope with the problems of illiteracy (or marginal literacy) in the Armed Services. Two of these strategies, selection and classification, depend for their success upon adequate assessment instruments for predicting who will be successful or unsuccessful on the job. The remaining two strategies involve the development of training programs for lower aptitude, less literate personnel. In one approach, the person's low literacy skills are considered as a fixed characteristic of the person, and modification of job technical training programs to adapt to the person's lower-than-normal capabilities is attempted. In the fourth strategy, an attempt is made to produce a more adaptive person through the provision of basic literacy skills training.

Nonacceptance of Illiterates

One way to overcome the problems resulting from low literacy skills in inductees is to avoid them. From time to time, the Services have raised admittance standards with the result that large numbers of marginally literate men were excluded from service. In general, standards have been raised during intervals of relative military quiescence and lowered during periods of military activity, such as the Korean and Vietnam conflicts.

There are, however, several problems associated with the strategy of nonacceptance that limit its fruitfulness. For one thing, as with many other abilities, it is not a simple matter to accurately assess a man's literacy skills at the selection station. Hence, large numbers of potentially useful men may be turned away, while some who are not useful may be accepted.

The problem of accurately selecting men on the basis of their literacy skill is compounded by the fact that, until the research reported by Caylor in this volume, no attempt had been made to accurately identify literacy skill levels required by military jobs and training schools. Therefore, there were no adequate criteria for selecting cutoff points on selection tests.

Both of the foregoing problems are concerned with assessment — the first with assessing the man, and the second with assessing the job requirements. A third problem affecting the usefulness and desirability of the strategy of nonacceptance concerns training. Manpower needs are such that it may become necessary, under conditions of large-scale mobilization, to enlist marginally literate men. If these men are not accepted during peacetime and the training methods needed to keep pace with technological change developed to effectively train such men, then new training techniques and methods will have to be developed under the stress of mobilization, when expediency rather than effectiveness may be the predominant training motive.

Perhaps the most significant results of the nonacceptance into the military of marginally literate people is that a large and needy segment of the population is not able to reap the benefits of the training, education, social development, and practical experience that accompany military service.

Limited Assignments

The problem of using marginal literates also has been dealt with by assigning these people to jobs that have "minimal" requirements for reading. This strategy has not worked well for several reasons. First, as with the other strategies reviewed, adequate definitions of the requirements for reading skills in different jobs have not been available; hence, it has not been possible to accurately state "minimal" requirements for reading skills.

A second difficulty is concerned with selecting job proficiency levels for establishing reading requirements. Individuals qualified for entry-level jobs are not necessarily qualified for advanced level jobs. This may be a particular problem for marginal readers because higher level positions tend to place higher demands on reading skills. In work for the Navy, Sticht et al. (1976) found that the amount of time spent in job-related reading increased as a function of rank, with pay grades E1-E3 reading 0.7 hours; E4-E5 reading 1.7 hours; and E6-E7 reading 2.1 hours per day. Similar findings were reported in a Department of the Army (1968) survey of the primary reasons for reading among Army officers and enlisted men. For enlisted men, the percentage of job-related reading gradually increased from 2% for E1s, to 25% for E8s, and then dropped to 16% for E9s. Thus, higher levels of pay and responsibility seem to require a greater amount of reading. To assign personnel to nonreading "tracks" at the start of their careers is to place a definite limitation on their career development and restrict their range of utility (say, through cross-training or re-training as new jobs open up) to the Services.

A third problem with the limited assignment approach to dealing with personnel of low literacy skills is that in most instances, advanced-level positions are filled with personnel from the entry-level jobs. The assignment of a man with the marginal requirements needed for an entry-level job may result, perhaps because of combat casualties, in his promotion to a leadership position, with possible devastating effects for him and the men he leads.

Another difficulty with the policy of assigning the marginally literate to a job having relatively low requirements for literacy and arithmetic skills concerns the overall effectiveness of that entire job field. This may be reduced if the job becomes flooded with marginals. Therefore, some means are necessary for distributing these people equitably among jobs.

Methodological Considerations in Determining Reading Demands of Jobs

The strategies of nonacceptance and limited assignments for marginally literate personnel both require, for their most effective implementations, information about the reading demands of jobs. Questions such as "What are the reading requirements of different jobs?" "How well does a person have to read to be able to do the job?" "What jobs can less literate personnel perform?" all require, for their answers, information about the reading demands of jobs. However, as the following review reveals, the determination of the reading demands of jobs is not a straightforward matter.

Review of literature indicates that reading demands of civilian and military jobs have been stated in two ways. In one way, reading demands are stated in terms of the types of reading materials and reading tasks the job involves. In a second approach, reading demands of jobs have been expressed as a single index number — the reading grade level of general reading ability needed to perform well on the job. The first approach can be called the *summary task statement* method of stating reading demands of jobs. The second approach can be called the *summary index number* method of stating reading demands of jobs.

Estimates of Job Reading Demands Given in Summary Task Statements

In this approach, job analysts or others usually interview management personnel, and sometimes workers, to determine whether or not a given job requires the use of reading materials. If so, then a simple statement to that effect is recorded in the job requirements.

As an example of this approach, we can cite the Army Regulation, *Enlisted Military Occupational Specialties* (AR Reg 611-201, 5 Jan 67). In this regulation, reading requirements for the Field Radio Repairman are given as: "Requires verbal ability to read and understand technical material pertaining to maintenance of field radio equipment." The mechanic (Ground Vehicle Repairman): "Requires verbal and reasoning ability to read and understand technical material pertaining to equipment being maintained . . . " The Military Policeman: "Requires verbal ability to interview witnesses and interrogate suspects, prepare written reports of findings, and read communications."

Recent work sponsored by the Navy (Sticht et al., 1976) describes another method of specifying reading tasks. In this case, a checklist of materials is compiled and people indicate which kinds of materials they read in their job; e.g., notes, notices, messages, manuals, etc. A similar method was used by Sharon (1972) to obtain an indication of what kinds of materials the adult population in the Unites States read. While such methods more precisely identify *what* people read than does the foregoing approach, they fail to identify *why* people read certain materials.

A recent project by the Department of Manpower and Immigration of Saskatchewan, Canada, has refined the method of summary task statements to obtain a more finely grained view of the reading tasks performed in various

career fields (Smith, 1975). In this study, an attempt was made to find out both *what kinds* of materials are read in these jobs (e.g., notes, memos, letters, directions, instructions, policy manuals, procedural manuals) and *what reading tasks* are performed in these jobs (e.g., read to locate and understand the main point or idea, to follow directions, to put details in sequence or order, to notice and interpret how facts or details are related, to make comparisons).

To obtain this information, interviewers at times showed pictures of the general type of material they were talking about. For instance, in determining if a given job required the reading of graphs, two graphs were shown as exemplars and interviewees were asked to indicate whether or not they read similar graphs in performing their job.

Thus, this method of deriving summary task statements goes beyond the preceding method in that it deals not only with *what* must be read in greater detail, but on *why* a person reads in a given job. Additionally, the Canadian work has distinguished entry-level reading requirements from advanced level requirements by asking respondents to indicate what reading tasks they performed when they first entered the job, and what reading tasks they presently perform.

It is clear that the Canadian method provides more useful information than the typical summary task statement or list of materials people read. However, even though the Canadian approach far surpasses the typical summary task statement in specification of reading tasks, it still fails to precisely specify reading tasks. For instance, the question of *what materials* job holders read versus *why* they read are asked separately, without reference to one another. Thus, though we learn that, in a given job a person reads policy manuals, and this same person reads to follow directions, we do not know that the person reads to follow directions in policy manuals. A more explicit statement of a reading task would include both a statement of what was read and why it was read.

There is also reason to question whether people can rate their performance of a generic reading task, when all we can present are specific displays (texts, figures, tables) with their specific content and (unspecified) dimensions of complexity, legibility, and so forth. The Canadian study presented numerous displays which people used as generic displays, even though each display was a "species" of the "genus". But that study did not have any way of determining the extent to which a genus versus a species basis accounted for people's judgments. The work by Sticht et al. (1976) found evidence to suggest that people may not respond to the genus aspect of a species display, but rather, to the content of the latter.

Additionally, responding to information displays with a simple "yes, I read things like that" or "no, I do not read things like that" does not define a *task*, because to "read" is not defined; though, since figures and tables are considered to be "read", a general definition like "extracting information from visual displays" may be necessary. But, extraction of information can go on at various levels. For instance, one can extract

information about the kind of type (pica?) used; about the color of the ink, etc. Or, one can extract information useful for constructing ideas represented by the message encoded in the printed display. If the latter is the goal, then we need to know what type of information is being sought for what types of ideas and with what type of given display.

Estimates of Job Reading Demands Given in Summary Index Numbers

Educator Estimates of Job Reading Demands: Perhaps the most widely used method for stating reading demands of jobs is the reading grade level (RGL) index number. This appears to have come about because the need to state reading demands of jobs has usually occurred along with the need for literacy training programs. For instance, as Hiller indicates in this volume, in World War II, the Army was required to establish literacy training for many illiterate personnel. It was decided that the goal of such training should be to produce skill in reading up to the level typically achieved by children in the fourth grade. This level was established by educators as a "guestimate" of the minimal reading demands of Army life. Since WW II, whenever large mobilization efforts have resulted in the induction of large numbers of marginally literate personnel, such as with Project 100,000 in the mid-sixties, the Air Force, Army, and Navy have all established reading training programs having an RGL index number as the goal of the program (cf, the papers of Part II of this volume). In Project 100,000, the Air Force established a goal of grade 6 reading ability for its reading program, the Army set its goal at a grade 5 level, and the Navy produced graduates in the 5.0 to 5.5 RGL range. No documentation of the bases for these estimates has been found.

Job Reading Demands Derived from Job Analysts' Judgments: In civilian settings, the Department of Labor (DOL) has sought ways to establish the reading demands of jobs, again for the purpose of providing guidance on objectives for programs for preparing marginally literate persons with marketable reading skills.

In the DOL approach, job analysts estimate the levels of General Educational Development (GED) required for various jobs, based on interviews with job incumbents, supervisors, and observation of the job being performed. Jobs are then categorized as requiring one of six levels of GED. These levels have been developed to roughly parallel school-based educational development. Hence, for example, a GED of level 1, is said to approximate the education obtained in grades 1 through 3, level 2 parallels 4 through 6th grade education, etc. (Phillips, 1970). Thus, when a job is assigned a GED level, it has also been assigned a reading grade level. To say a job has a level 2 GED, is to say it requires a 4th- to 6th-grade reading ability.

This approach, like that of the educator's estimate, is a *judgmental* approach, which calls for an estimate by the job analyst. Though relatively low in cost, the lack of specificity in the rules for arriving at a judgment of the GED level, and hence reading level of the job, and the absence of empirical information which validate the estimates of the job analyst (or educator) renders this method too imprecise and uncertain for establishing reading demands of jobs.

Readability Estimates of Job Reading Demands: The development of methods for estimating the reading grade level of difficulty of various reading materials, by specially constructed readability formulas, makes possible a relatively low cost method for estimating reading demands of jobs. By applying a readability formula to samples of job reading materials, an average reading grade level of difficulty for the materials can be computed, and used to represent the reading requirements of the job or job training program.

Readability formulas have typically been constructed by two means. In one approach, prose passages from school textbooks of various grade levels are sampled and features such as average sentence length, number of one-syllable words in 100 words, and so forth, are determined. These features are then used in correlational analyses to find out how well they can be used to predict the school grade level of material. Generally, one finds that average sentence length and word length increases as materials from higher grades are sampled. Because of this positive correlation, it is possible to obtain a sample of material, determine average values for features such as sentence and word length, enter these values in a regression equation, and then state that materials having those values are typically found, say, in the 6th grade of school. Therefore, the material is said to be of 6th-grade difficulty, and by definition, requires 6th-grade reading skills.

The preceding approach to the construction of readability formulas does not involve any direct measure of people's abilities to comprehend the materials. It is not known in that case whether 1/3, 1/2, 1/4, etc. of the 6th-grade students can actually comprehend the material having the structural features typical of materials found at the 6th grade. For this reason, most readability formulas have been constructed to relate features of textual material such as sentence and word length to performance on tests of comprehension of the material. Then a reading grade level is assigned to the material by setting a criterion of accuracy on the comprehension test, say 70% correct, and determining the earliest grade level at which some designated proportion of people, say 50%, get 70% correct on the comprehension test. If it happened that, on a given comprehension test, it was not until the 6th grade that 50% of the students got 70% correct, then that material would be assigned a reading difficulty or readability value of 6th grade.

To use such a formula for determining the reading demands of a job, one must:

1. Identify job reading materials.
2. Sample the materials representatively.
3. Calculate the critical features (e.g., average sentence length).
4. Use these features in the readability formula to obtain an estimate of the comprehension score 70% of the people would get if they were to take a comprehension test like the one used to construct the formula.
5. Convert that score to a grade level score using appropriate tables. (If step 4 provides a direct estimate of reading difficulty in reading grade score equivalents, then step 5 is unnecessary.)
6. Obtain the average reading grade level of all the materials sampled. This is the reading level demanded by the job.

Clearly, the foregoing method for determining reading demands of jobs is a function not only of the features of the text materials, but also of the comprehension test items, the criterion set as acceptable on the comprehension test (e.g., 70% correct), and the criterion set for the proportion of people at each grade level (e.g., 50%) who must achieve the criterion score. As an example, in a study for the Army (Caylor, Sticht, Fox, & Ford, 1973) changing the criterion score on the comprehension test from 30% correct to 35% correct, a 5% increase, changed the reading grade level assigned to some materials by as much as 2 to 3 grade levels!

Beyond these problems inherent in the methodology for developing readability formulas, there are additional problems in using this approach for determining reading requirements of jobs. For one thing, it may not be possible to obtain a representative sample of job materials, nor even to determine the proper domain of materials from which sampling should be done. In this regard, a major difficulty can arise due to the distinction between the *formal job task specifications* and the actual, *informal job tasks* which are performed on a day-to-day basis. The problem is that, if supervisors, management, or content experts are consulted to find out what reading materials a person must use in doing the job, they are likely to base their statements on their conception of the *formal*, or even idealized, job and prepare a list of materials which no one could reasonably be expected to encompass, and which are in fact not used in the work-day activities.

Interviews with employees may also produce a distorted sample of job reading materials if employees tend to "fake good" in response to how much and what it is they read. On the other hand, experienced employees may fail to report certain reading materials which they used in first entering the job, but no longer use. It is necessary, therefore, to ensure that new and old employees are interviewed in identifying job reading materials.

A second problem with readability measures is that they tend to set reading requirements somewhat higher than do the other empirical methods to be discussed below. Work by Caylor, et al. (1973) showed that some Army jobs would require 12th-grade or higher reading skills if readability factors alone were considered. But because many persons with reading skills well below the 12th-grade level were successfully completing career education programs and performing successfully on the job, we must regard the reading requirements suggested by the readability analysis with some caution.

Estimates of Reading Demands of Jobs Stated in Index Numbers Derived from Correlating Reading and Job Proficiency Measures: A general method for estimating job reading requirements is the traditional psychometric procedure used for validating selection and classification tests. In this procedure, performance on a reading predictor test is related via correlational techniques to performance on a job proficiency test. If a sufficiently high relationship exists, cut-off scores on the reading predictor variable can be selected to increase the probability of obtaining students or employees who will reach an acceptable level of achievement or the job proficiency criterion measures.

In research for the U.S. Army (Caylor, this volume) this psychometric model was applied to determine reading demands of four jobs: Cook, Mechanic, Supply Clerk, and Armor Crewman. Some 400 men experienced in each job were administered standardized reading tests and two measures of job proficiency: a 4- to 5-hour individually administered job sample test, in which men performed actual job tasks derived from extensive job and task analyses; and a job knowledge, paper-and-pencil test designed to include questions about information actually needed to do the jobs.

To establish reading demands of these jobs, it was necessary to develop decision rules for determining what level of reading skill was associated with a desirable level of proficiency in the job sample and job knowledge tests. The decision rule finally arrived at stated that the lowest level of reading which should be used to establish the reading demands of the job is the level at which men would not be expected to be overrepresented in the bottom quartile of performers. In this case, the criterion of job proficiency was a relative one, being defined in terms of quartiles. It was judged that if a person was performing at a level below 75% of his fellow job incumbents, and if this was systematically related to reading skill, then it would not be overly demanding to target the objectives of a reading training program to a level associated with not being overrepresented in the lowest one-fourth of job performers.

It should be noted that, in the foregoing case, the *purpose* of the exercise in estimating the reading demands of jobs — in this case to derive a goal for reading training — entered into the formulation of the decision rules for setting criterion levels of job proficiency and for relating reading skills to those criteria on both job sample and job knowledge tests. In the project under discussion, all persons tested were, in fact, working in their jobs and, by virtue of this fact, could be construed

as successful job performers. Hence, the decision rules had to be formulated to discriminate among successful job performers for the purpose of deriving goals for reading training. If the purpose of the determination of reading skills demanded of a job is to screen out people or classify them for limited assignments, then a different criterion (e.g., attrition) and different decision rules would be called for.

While the method of correlating reading and job proficiency criteria is a general approach with established methodological techniques, it has some serious limitations. For instance, because the job proficiency measures are likely to be only indirectly mediated by knowledge which may have been previously learned by reading, relationships of reading to job sample performance should be expected to be smaller than relationships among general reading measures and job reading tasks or other paper-and-pencil measures of job proficiency. These considerations, plus the fact that the costs of constructing and administering an extensive job-sample test to job incumbents in a representative sample of jobs are prohibitively high, would seem to mitigate against the use of job sample criterion tests for all but fundamental research purposes.

Correlating reading skill with job *knowledge* is a standard, straightforward approach to determining job reading requirements which might readily be used in various job or training settings with existing personnel job knowledge tests. However, it may not be certain that reading test scores are correlating with job knowledge per se — they may be correlating with the reading demands of the job knowledge test. Application of this method on a broad scale would seem to be justified only if it could be shown that job knowledge, and not general reading skill per se, was needed for scoring well on the job knowledge tests.

Furthermore, the method of correlating reading skills with job proficiency measures such as job sample tests, job knowledge tests, attrition, and similar criteria to establish reading levels of jobs provides no direct indication of how well a person must read to perform job *reading* tasks. Many job tasks can be performed without reading, and may be learned by watching and imitating others. For this reason, reading demands of jobs may be over- or under-estimated whenever non-reading criteria are used to establish relationships of reading skills to job proficiency.

Estimates of Reading Demands of Jobs Derived from Job Reading Task Tests: To render the correlational analysis technique more directly relevant to job *reading* demands, it may be possible to relate general reading skill to performance *not* of job tasks in general, but rather to performance of job *reading* tasks; i.e., tasks in which reading is required.

In research of this nature for the Army (Sticht, 1975c) Cooks, Mechanics, and Supply Clerks were interviewed at their job sites. In the interview, each man was asked to identify reading materials he had used in performing some job task. Copies of these materials were obtained and analyzed as to the reading tasks involved in using them. Tasks such as

"reading tables of contents", "reading indexes", "reading procedural directions", "reading tables of standards and specifications" were identified. Special job reading task tests (JRTT) were constructed which tested the ability of men to perform the different job reading tasks.

The JRTT and a general reading test were administered to several hundred Army personnel. With these two sets of scores, it was then possible to relate various criterion levels of achievement on the JRTT to the general reading grade level of ability needed to achieve this criterion.

Using this approach, and by averaging over the various reading tasks, it is possible to indicate general reading levels associated with various criterion levels of performance on the job reading tasks as a group. Given this information, and a decision about the criterion level of performance which job performers or job aspirants should display on the job reading tasks, a general literacy requirement can be estimated for each job.

In the Army research, it was found that if a criterion of excellence was chosen such that 70% of the people were expected to get 70% correct on the job reading tests, the general reading requirements for the Cooks field was 7th-grade level, for Mechanics - 8th-grade level, and for Supply Clerks - 12-grade level.

The *job reading task test method* represents the most direct approach to determining job reading requirements in that it takes as its criterion measure the reading score on the JRTT, a sample of actual and commonly used job reading materials. To the extent that the job reading passages constituting the JRTT comprise or represent all the reading tasks of the job, and to the extent that the tasks people are asked to perform on the JRTT represent tasks people have to perform on the job with those materials, then the ability to read the JRTT passages <u>is</u> the ability to perform the job reading tasks and thus, to meet the job reading requirement.

Additionally, it should be noted that using the JRTT method, reading skill level requirements will change depending upon the criterion of performance selected. This latter problem, i.e., the problem of specifying a criterion, must be dealt with in any approach to the determination of reading requirements in which criterion performance measures are obtained. The question is one of "how good is good enough"? It is possible to say that all people should be able to perform all reading tasks with 100% mastery. But if there are restricted manpower pools, and if many job reading tasks are quite complex, this would seem an unrealistic goal.

<u>The Validity of Estimates of Reading Demands of Jobs</u>

From the foregoing discussion it should be apparent that there is no such empirical "thing" or "stuff" or "event" or "condition" known as "*the* reading demands of a job". Reading demands are not *discovered*, they are *created* by procedures which are more or less systematic and performed according to more or less specifiable rules. The question of the *validity*

of any estimate can only be answered with respect to a model or theory of job-related reading which would define systematic procedures for obtaining estimates of reading demands of jobs for various purposes permitted by the theoretical constructs involved.

In the absence of such a theory, it may be easier to know that an approach is not valid than to know that one is. For instance, the question of validity seems clear with the readability approach: it in no way involves figures and tables in the estimate; and Sticht, et al. (1976) found in their structured interviews with Navy personnel that only 32% of the reported reading tasks involved texts only; 62% used figures or figures and texts combined (the remaining 6% was "no response").

The job reading task test method demands a theory for constructing reading tests which can be applied to a sample of reading tasks. But, if we are to be certain that these reading tests match the information processing which was done in the actual performance of the reading task on the job, then the theoretical constructs in the "reading test model" must be incorporated into the interviews or other approaches used to obtain the sample of reading tasks, because, since reading is a "private" act, it is necessary to query people about their information processing during the performance of a reading task in order to discover whether or not they perform the types of information processing involved in the model of reading.

Perhaps the most valid estimate of the reading difficulty level of military jobs would be obtained by having personnel of various reading levels perform the actual job tasks and reading subtasks which were reported to have been performed. But, what would be scored? Completion of the job task? Perhaps some could be completed even though some or all of the reading material was not understood. How would one know? Would speed of reading be scored? Would it matter? Could one then develop an inventory, as in the Canadian study, for general use in determining reading demands of all jobs which job analysts could use in a discriminating, reliable manner? How would this be determined?

These considerations indicate some, though clearly not all, of the conceptual and methodological problems involved in the use of the strategies of nonacceptance and limited assignments based on reading requirements of jobs for coping with problems engendered by low literacy skills of personnel. Hopefully, awareness of the procedures and problems discussed here will be useful in future research to develop more adequate techniques for the assessment of the reading demands of military jobs. Additional discussion of these types of problems can be found in Bormuth (1975).

Modification of Training and Job Requirements

A third strategy for dealing with the problems of low literacy skills is to redesign training and job materials to minimize the importance of such skills. Under this strategy, training schedules and practices may be modified to meet the skill levels of different individuals. Thus, written instructional material may prove adequate for certain individuals, but the same instructions might best be presented in some other way for individuals having relatively low reading skill levels.

The application of this approach is illustrated by consideration of two primary recommendations regarding the training of Mental Category IV personnel presented in a review of some five years of study by the Naval Personnel and Training Research Laboratory.

> 1. Do not assign Group IV personnel to Navy ratings that have high verbal, computational, or conceptual requirements. (The storekeeper rating is an example.)
>
> 2. In general, training programs should minimize reading requirements (including written tests) and theoretical knowledge and should utilize pictorial materials, simplified terminology, and practical illustrations. The actual performance of job skills should be stressed.
> (Steineman, Hooprich, Archibald, & VanMatre, 1971, p. 8.)

The Navy laboratories are not alone in their conclusions regarding the advisability of using reading materials for training Category IVs. In a report of an extensive series of studies on the development of training techniques for use with Category IV personnel, Army research concluded:

> For the low aptitude groups, printed programs or programmed texts were the least effective technique. On the basis of this study we conclude that written material as the primary medium should be used with low Category IV personnel only as a final recourse. (Bialek, Taylor, & Hauke, 1973, p. vii.)

Papers in the present volume present evidence of the concern for reading demands of jobs and job training programs by reducing the reading difficulty level of materials. (See papers by Burkett, Curran, Hiller, Johnson, Kern, and Muller.) As these papers suggest, lowering the difficulty level of the printed materials is an important step. However, this approach, by itself, cannot be expected to "solve" the literacy problem. The extent to which the reading difficulty level of technical materials can be lowered is limited. After a certain point, it becomes necessary to start simplifying and deleting the more demanding job knowledge requirements. The result can be to render a person useable only in a very limited job assignment. This may be a satisfactory state of affairs when manpower is unlimited and jobs can be fractionated and specialists who do nothing but a small part of the job can be used. However, we are in a time in which weapons systems and associated publications systems are growing more complex and voluminous, as Muller's paper in this volume reports. Accompanying this growth in weapons/documents systems has been a

reduction in the numbers of personnel in the Armed Services. The fact that there is more to do with fewer people precludes the "divide and conquer" strategy involving more and more specialists. In fact, what is needed are more and more generalists, and to have flexibility, a person must have the fundamental literacy/cognitive skills to adapt to new job demands as they occur.

On the other hand, as Glaser comments in this report, caution is required to assure that job training programs are not overly complicated by the excessive use of language — written or spoken. Olson's (1975) trenchant discussion of the problem of over-emphasizing literacy skills in training programs is particularly apt.

> The nature of schooling per se tends to make literacy relevant to many nonliterate tasks simply because the means of instruction, worked out to permit the group instruction of large numbers of students, tends to require literacy. Many tasks, if taught in a practical context, on a master-apprentice basis, allowing for attempts at performance accompanied by feedback, could be taught with very low literacy demands. To teach the same tasks on the basis of group instruction immediately transforms the task into a literate one — presumably this is the case whether or not one uses printed materials or lecture materials. In both cases, a practical task is changed primarily into a verbal one — calling for the analysis and memory of statements instead of a pattern of action. This is well illustrated in the familiar poem by Henry Reed, called Naming of Parts. The poem illustrates our problem. The use of a rifle by a soldier is certainly a practical action. But to permit group instruction in its use, and perhaps to permit group performances of certain kinds, the practical action is dissolved into dozens of verbal descriptions which are then memorized. In this way, practical action is transformed into a task of verbal learning. Once any practical action is so transformed, literacy skills become relevant to the acquisition of practical action. Hence, the reportedly high correlation between literacy skills (IQ roughly) and success in such things as military training. But that correlation occurs, it may be argued, simply because of the form of instruction employed; it may not occur with the practical actions themselves. . . . Here is Reed's poem:
>
> Today we have naming of parts. Yesterday,
> We had daily cleaning. And tomorrow
> morning
> We shall have what to do after firing. But
> today,
> Today we have naming of parts. Japonica
> Glistens like coral in all of the neighboring
> gardens,
> And today we have naming of parts.

This is the lower sling swivel. And this
Is the upper sling swivel, whose use you will
see
When you are given your slings. And this is
the piling swivel
Which in your case you have not got. The
branches
Hold in the gardens their shy, eloquent
gestures,
Which in our case we have not got.

This is the safety-catch, which is always
released
With an easy flick of the thumb. And please
do not let me
See anyone using his finger. You can do it
quite easy
If you have any strength in your thumb.
The blossoms
Are fragile and motionless, never letting
anyone see
Any of them using their finger.

And this you can see is the bolt. The purpose
of this
Is to open the breech, as you see. We can
slide it
Rapidly backwards and forwards; we call this
Easing the spring. And rapidly backwards and
forwards
The early bees are assaulting and fumbling
the flowers:
They call it easing the Spring.

They call it easing the spring: it is
perfectly easy
If you have any strength in your thumb: like
the bolt,
And the breach, and the cocking-piece, and
the point of balance,
Which in our case we have not got; and the
almond-blossom
Silent in all of the gardens and the bees
going backwards and forwards,
For today we have naming of parts. Henry Reed

There are two points to be made. First, the very demands of large group instruction tend to translate even practical tasks into highly literate ones. Such translation is prejudicial to people with lower degrees of literacy — such learners will make slow progress and will be regarded as dull, even if they ultimately make the best performers. (Olson, 1975, pp. 149-152.)

While there can be no denying that marginally literate personnel can be trained in motor skills using "hands-on" methods and utilized in less demanding occupations, there is reason to question the long-term value of the strategy of nonverbal training, both from the Services' perspective and that of the trainees. For one thing, though reading materials can be eliminated from training programs, it is doubtful that they can be removed from on-the-job use. The reason is that printed materials serve as repositories of needed facts (e.g., standards and specifications, procedural directions) which, if not accessible by means of reading skills, must be memorized in training. Since equipment and procedures may change *after* a man has completed training, it would prove very expensive to redesign a training program to provide nonreading, hands-on training every time such changes occur. A primary function of print is, in fact, to replace the need for direct, hands-on, personalized instruction, thus expanding the individual's range of utilization.

With regard to career development, the strategy of avoiding the use of printed materials, starting early in a man's career, may foretell a limited level of achievement later on when actual job demands call for greater use of reading materials, as indicated above.

The problem for the design of training is to insure that the training prepares personnel for performing all entry-level tasks — including both motor skill and literacy tasks.

Remedial Training

A fourth strategy sometimes used by the Services in an attempt to solve the problems associated with low literacy is to provide remedial training in reading. The chapters by Fletcher, Hiller, and Burkett review remedial reading training efforts in the Services up to the end of the '50s, while Caylor, Duffy, and Groff discuss current literacy training programs in the Army, Navy, and Air Force.

For the most part, literacy training in the Armed Services has been closely associated with major mobilization efforts; namely, World War II, the Korean Conflict, and Vietnam. During such efforts, large numbers of personnel are needed, and it becomes necessary to recruit large numbers of persons having low literacy skills. The typical response, then, has been to initiate a remedial reading program to train these people to read at some minimal level *presumed* adequate for coping with the basic demands for reading in the military. For the basic reading programs established during Project 100,000, the targeted levels for reading achievement ranged from grade 5.0 to 6.0 across the three Services.

Only a limited amount of data bearing on the effects of military literacy training programs on job proficiency have been found. Much of the research to evaluate literacy training conducted prior to Vietnam, summarized in detail in the book *Marginal Man and Military Service* (Department of the Army, 1965), failed to demonstrate benefits of literacy training on either job training or job performance, measured in various ways —

supervisor's ratings, retention, pay scale achievement, conduct, and others. However, these studies also failed to demonstrate that the various literacy training efforts improved reading skills over those of control groups — where such comparisons were made. If proficiency in reading is not genuinely improved, it is meaningless to search for the effects of such non-improvement on subsequent measures of job training or job performance proficiency.

Evaluations of literacy programs since Vietnam with respect to their impact on training or job proficiency are scarce, and in no case do they isolate the effects of literacy training per se from the effects of having an additional 4 to 8 weeks of time to adjust to military life and, in most cases, to receive additional training in military subject matter offered during the period of literacy training.

In the case of the Army Preparatory Program initiated at the outset of Project 100,000, Fisher (1971) compared men who were successful and unsuccessful in reading the fifth-grade level of reading and found that they did not differ greatly on most indices of military status and performance, though successful trainees were slightly more likely to achieve a higher pay grade and to be judged eligible for reenlistment.

In an unpublished report by Zaccaria (1971), graduates of the Air Force remedial reading program showed a much larger completion rate for basic military training than comparable (in reading) personnel who did not have remedial reading training. However, the remedial reading program also offered four hours of military training daily, so it is not certain that the reading training per se resulted in the different success rates in basic military training.

Duffy's paper in the present volume points out the inconclusive nature of data on the impact of the Marine Corps and Navy reading programs on later Navy effectiveness. Apparently, adequate studies of this relationship are not available.

Though it may seem a fairly straightforward problem, the evaluation of the effects of literacy training on job training and job performance proficiency is fraught with many difficulties. What should the measure(s) of job proficiency be? If hands-on job sample tests are used, is there any reason to believe that limited literacy training will have any effect on such performance measures? Work by the Human Resources Research Organization (Vineberg & Taylor, 1972) compared the job proficiency of men given job training (8 or so weeks of school training in the performance of job tasks, *not* reading training) as Cooks, Vehicle Repairmen, Supplymen, and Armor Crewmen to that of men given on-the-job training. The proficiency measures were job sample tests based upon task analyses of these jobs. Results indicated that, for Repairmen and Cooks there was a statistical (but not practical) difference in favor of the men given formal job training, while there was no statistical or practical differences between the formally trained and on-the-job trained Armor Crewmen and Supplymen. (In these comparisons, effects of AFQT, time on the job, education, and age were controlled by covariance techniques.)

If eight or more weeks of formal *job training* may have little or no effect on subsequent job sample test performance in these types of jobs, is there any reason to believe that four, six, or even 13 weeks of *reading* training will be reflected in such measures? Probably not. Many job tasks can and will be learned by "show-and-tell" — especially tasks of high frequency and criticality — i.e., the types of tasks emphasized in task analyses and typically used to construct job sample tests.

Furthermore, except for the *profoundly* illiterate (e.g., persons reading below the grade 2.0 level) which most of the military reading students are *not*, it seems unreasonable to expect four to six weeks of literacy training to meaningfully affect retention rates, rate of promotion, conduct (AWOLS, court-martials), paper-and-pencil job knowledge proficiency test scores, or supervisors' effectiveness ratings. For the *profoundly illiterate*, improvement on any of these indices would be contingent upon successful literacy training up to some requisite level. The Air Force report by Zaccaria (1971) suggests that it is difficult to effect much reading improvement with such personnel in as much as 13 weeks of reading training. In this study, 85% of the men who entered reading at the first-grade level exited reading at or below the fourth-grade level.

Perhaps the most reasonable criterion measure of job proficiency against which the effects of reading training might be evaluated would be measures of a person's ability to perform necessary job tasks involving reading. *Necessary* reading tasks could be defined as those reading tasks which a person would have to perform *if* he was working alone on the job and had no one to turn to for help in doing a job task. This is a somewhat artificial situation, since it is rare for personnel to work under such restrictions. Usually, the marginally literate person can rely upon someone more literate to handle the reading needed in various job duties. But if it is desired that each individual be able to perform *all* significant job tasks, including the reading tasks, then it seems reasonable to evaluate the effectiveness of literacy training by comparing the performance on job reading task tests of graduates of the literacy training program with that of persons of equal pre-literacy training reading skill levels. It seems reasonable to argue that if reading training does not improve a person's ability to perform significant job reading tasks, then it is unlikely that the reading training will affect other, less directly reading-related indices of job proficiency, such as retention rates, supervisors' evaluations, job sample tests, etc. On the other hand, it does not necessarily follow that improved competency in performing job reading tasks will be reflected in other, non-reading related indices of job proficiency. In fact, data from HumRRO (Vineberg, Sticht, Taylor, & Caylor, 1971) indicated that job knowledge and job sample test performance for the four jobs mentioned earlier correlated only about .5, while supervisors' ratings were only trivially correlated with either of the other indices of job proficiency. There is thus evidence to suggest that improvement on one index of job proficiency — e.g., the use of job reading materials — will not necessarily be reflected in improvements on other indices of job proficiency.

To date then, the strategy of dealing with literacy problems in the Armed Services by providing brief remedial reading programs has not been demonstrated to be an effective approach to developing more competent, effective readers capable of contending with the reading demands of technical training programs and many reading tasks demanded by military jobs. Usually, only one or two years of gain in "general" reading proficiency are produced, which may typically shrink to less than one year after a few weeks following the reading training (see the papers by Duffy and Caylor).

It seems likely that the belief that adults can be taught to read in brief periods of time is due to our not having adequately understood what we mean by the concept of "literacy". A major (though not the only) problem is the tendency to confuse the information processing skills of reading and writing with the knowledge expressed or received by those skills. For example, in a reading vocabulary test, it is possible to miss one item because one lacks knowledge of the word meaning. A person might be perfectly capable of saying the word out loud, however, thereby indicating skill in converting the printed word into a spoken word.

It is clear that if a person does not know the meaning of a word, then our instructional practice must differ from what we would do if the person knew the meaning of a word when that word was spoken, but could not recognize the word when presented in the printed form. If a person does not know the meaning of a word when presented in the printed form, and also does not know the meaning of a word when presented in the spoken form, then we may conclude that the person will have difficulty understanding reading material containing that word, but the problem will not be a reading problem. Rather, it is a language problem which no amount of instruction in sight/sound correspondences or sight/word recognition will remedy.

This distinction between the oral language knowledge a person has and skill in recognizing the same language form in print is a crucial one for understanding why brief reading programs are likely to produce only slight improvements in the ability to perform various tasks involving reading. The unspoken assumption of reading training programs for adults is that many illiterate or marginally literate people have a fairly well developed oral language capability. Hence, a relatively brief period of training in decoding skills, i.e., providing knowledge of sight/sound correspondences and opportunities for practicing this knowledge to develop skill, is considered sufficient for unlocking the person's oral language skill via the printed word. At the extreme, we might imagine a person whose oral language skills/knowledge were equivalent to those of a typical 18 year old with 12 years of education, but who had simply neglected to learn to read. In this case, a relatively brief period spent learning the heuristics of decoding would permit this person to bring this fully developed language skill to bear on the printed word. In such a case, the person might gain as much as 12 years in reading skill with only a few weeks of decoding training!

A problem for adult literacy programs is that clients are likely to be poorly developed in *both* oral and written language skills. In Air Force sponsored research (Sticht & Beck, 1976) men in a military literacy training program were administered a test of auding[1] and reading comprehension. The results showed that both oral and written language skills were at the 5th-grade level. These findings are confirmed in research reported herein by Duffy using Navy personnel.

The fact that most marginally literate personnel score several years below the norm for their age group in auding indicates that their "literacy" problem is not restricted to the printed page. Rather, it reflects in large measure a language processing/comprehension problem. Low oral vocabulary scores indicate a lack of language knowledge, not just low skill in processing the connected discourse used in assessing paragraph comprehension. Because language knowledge is low, reading comprehension will tend to be low. Hence, even if a marginally literate person could decode printed language as well as he could spoken language, his low oral language skill would retard comprehension. An implication of this is that major improvements in reading skills of marginally literate personnel will require major improvements in language competence (e.g., vocabulary knowledge).

Carroll (1971) discusses problems involved in measuring growth in vocabulary (by auding or reading). He cites work by Edgar Dale in which Dale estimates that children may finish first grade with a vocabulary of 3,000 words. Then he estimates that they will add about 1,000 words per year from then on through high school, with high school seniors knowing about 15,000 words.

Accepting this rate of growth, if marginally literate personnel typically have vocabularies like beginning 5th graders, then they know about 6,000 words. To achieve a 6th-grade vocabulary, about 1,000 words will have to be learned. In a 6-week program in which six hours of active learning occurs each day (an unlikely situation) new words would have to be learned *and retained* at a rate of six per hour. To reach a 7th-grade level vocabulary, 12 words per hour would have to be learned and retained. The problem enlarges when it is recognized that this rate of learning is based upon typical learning rates of typical children. Marginally literate personnel, however, usually require anywhere from two to four times as much time to learn verbal concepts as typical recruits do (Fox, Taylor, & Caylor, 1969). For these people, it seems unlikely that they could learn and retain six to 12 words per hour for six hours a day, or even half this number. As Carrol has pointed out ". . . basic linguistic competence (at least with respect to grammar and vocabulary) is probably relatively unsusceptible to improvement except over long periods of time and with tremendous efforts . . ." (p. 130).

---

[1] Auding refers to listening to speech in order to comprehend and is a parallel term to reading, which refers to looking at printed language in order to comprehend.

The data regarding the effectiveness of current literacy training programs reported in Part II of this volume, and the foregoing considerations strangely suggest that the brief literacy programs the Services have operated in the past and are currently operating take only the first few steps of a long journey. Furthermore, it does not seem unreasonable to expect that even the progress gained in the current programs will be lost because no systematic efforts are made to continue to stimulate the less literate man to continue his reading development after completion of literacy training; no job-related, developmental reading materials are provided, and no officially scheduled time for continued literacy development is provided.

A similar concern for continued education and training in reading following the initial short-term program has been expressed by Kent et al. (1971) as a result of their examination of civilian job-related adult basic education (ABE) programs.

> Available figures indicate that a single ABE program by itself usually accomplishes very little. In most programs, the advances which occur in reading, writing, and arithmetic are somewhere around one grade level, or occasionally two grades. Advances in self-confidence may or may not be more substantial — no figures are available. But new self-confidence, even if present, is probably fragile.
>
> So the ABE graduate is unlikely to be in a markedly different employability situation than faced him before ABE. (Note the significance for ABE program goal definition — that program goal statements should refer to making a small start in an important direction, and that programs should be evaluated not so much by the amount of the start, as by the realistically useable momentum imparted in the right direction and by their working linkages with post-ABE activities.) Post-ABE help is required to capitalize on these small but important gains. Either the ABE program should have skills training and job placement or advancement built into it, or (more to the point) ABE should be built into a total employment or advancement program. If the ABE is realistically related to what is going to happen next in work or training, then its results will be reinforced and expanded upon naturally by the new work or training situation.
>
> Evaluation of a well organized total job-related ABE program will result in recommendations for improvements, perhaps to include additional time with ABE content at later dates (say 3 months, 6 months, or a year after "graduation"). These later sessions might be even more thoroughly integrated with job training than the original sessions were.

Here we are pointing out the importance of a context for ABE which will link it with skills training and work. This context can hardly be supplied by ABE itself, but it could be supplied by skills training or by work or by some combination. Many unions and employers have ongoing training programs, including classroom work and OJT, to which ABE could reasonably be adapted and which could supply this ongoing post-ABE support. Also, many schools (adult vocational schools, community colleges, and technical institutes) could (if closely enough related to employment situations) provide support. Any of these types of post-ABE support could also be made a part of the "selling-package" for ABE to increase its motivational attractiveness to students.

ABE is only a beginning, and much more beyond ABE is needed. But for many ABE students, a long-range view is entirely uncustomary. The ABE program should be set up so that it can lead on and on, but it may be inadvisable to burden the student with too much thinking about the future. For motivation and symbolically, glimpses of a distant future may help. But detailed learning activities should usually be based on more immediate student interests."

(Kent, Bishop, Byrnes, Frankel, Herzog, & Griffith, 1971, pp 94-95.)

RESEARCH TOWARD A COMPREHENSIVE CAREER-ORIENTED LITERACY TRAINING SYSTEM FOR THE ARMED SERVICES

Because of the many conceptual and methodological problems mentioned above, and the increasing recognition that the current, brief, one-shot remedial literacy training programs are inadequate for the task at hand, the Air Force, Army, and Navy research papers reported herein by Groff, Caylor, and Duffy each call for research which would study methods for validly characterizing literacy demands of jobs, and for the design and development of literacy training programs which systematically teach job-related vocabulary and reading tasks, and which are extensive enough to permit a long-term period for the development of literacy skills.

The interest in teaching job-related vocabulary and reading tasks stems from the desire to have personnel rapidly acquire the language needed for their military work. Since "general literacy" is made up of specific vocabulary, the teaching of job-specific vocabulary adds to a person's "general literacy" competence. However, the teaching of "general literacy" vocabulary, in a non-job-focused manner, will not necessarily include the various technical words and concepts found in military training programs and jobs. For these reasons, there seems to be a general consensus among military educational researchers that military literacy training ought to be focused on job-related reading requirements (see the papers by Caylor, Duffy, and Groff). This will both improve a person's "general literacy" ability, and permit the more direct transfer of skills from the literacy program to the job situation.

Recently completed research by Sticht et al. (1976) for the Navy suggests the possibility that basic military training, job technical school training, and correspondence course materials may be modifiable such that literacy skills and job skills training can be offered in an integrated manner. The concept is based on findings which suggest that much, perhaps 25% to 50% of what is currently taught in these training programs, is not considered relevant to their work by Navy personnel. That is, much of the current training programs may consist of "nice-to-know" but not "need-to-know" information. Thus, some instruction time might be saved if the training programs were systematically redesigned to effect a better match between training and job requirements.

Additional time savings might be accomplished by converting training programs from lock-step, lecture-oriented classrooms, to individualized, self-paced programs with flexible entry and exit. Where this has been done elsewhere in military training programs, considerable overall savings in man-hours (and hence cost) of training have been affected.

Because of the time, materials, and cost savings which may result from the careful systems engineering of training, and the move to self-pacing of instruction, it should be possible to incorporate literacy training into the programs for those who need it. While they would necessarily require more time to complete training programs under this system, the expectation is that overall training costs will drop, as those capable of accelerated programs are permitted to move rapidly through the streamlined programs.

Figure 1 presents a general plan for research and development to produce a comprehensive, integrated job skills and literacy skills career development system. The first phase of the plan calls for a thorough analysis of reading within the Services, including the determination of general reading levels for performing job-related reading tasks so that reading training objectives may be targeted to job requirements.

An additional aspect of the Phase 1.0 activities given in Figure 1 is the study of job skills training and the general education system within the Services to determine the feasibility of improving linkages in these systems, and potential for revising these systems to produce a more integrated job skills and literacy skills training system. The report by Sticht et al. (1976) illustrates the type of approach envisaged under the Phase 1.2 activities and provides information concerning the Navy's current career development system, including information regarding job skills training, career counseling, and general educational systems. A general finding is that the linkages among these subsystems are very loose — if not totally absent. Initial efforts toward a comprehensive career development system which includes integrated job skills and reading skills training should include the formation of tighter administrative and contentive linkages among these subsystems.

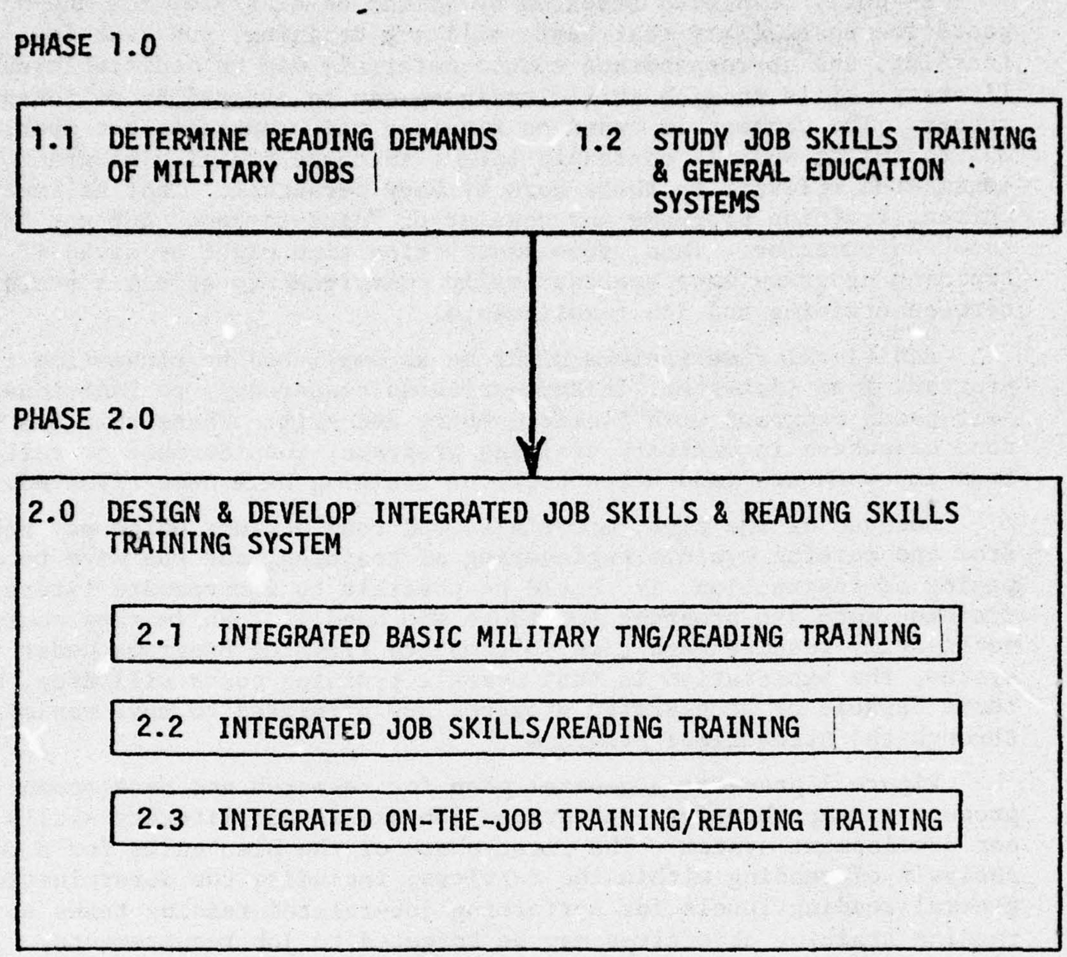


Figure 1. General Plan For Developing An Integrated Job Skills & Reading Skills Training System.

The Phase 2.0 activities of Figure 1 call for the evolutionary development of integrated job skills and literacy skills training programs for the three major stages of military career development: entry into the Service with basic military training; job technical school training; and assignment to on-the-job training at an active duty station. By providing opportunities for literacy training at each of the major career development stages, ample time is available for accomplishing genuine improvements in literacy skills.

RESEARCH AND THE TRANSFORMER: A PROGRAM FOR IMPROVING TEXTS

Michael Macdonald-Ross

The Open University, Institute of Educational Technology

In a modern industrial society such as ours people often get the information they need from documentary sources — pieces of paper or computer files. People cannot always meet face to face, and there can be quite a distance between an expert source and the user or client. It makes sense for us to spend time and effort on the business of communication. The task we face is simply this: *how can we organise complex information for the benefit of the reader* ?

When we see one person talking to another, both participants employ complex techniques so as to make themselves understood. These techniques of interaction may have taken millions of years to evolve, but are deployed instantly without self-conscious thought. It is much more difficult to make oneself clear to a reader who cannot see and who cannot ask questions. So, as one might expect, *communication breakdowns* are quite common, and can have the most serious effect. In England we have been suffering from what is known as Dutch Elm Disease — a fungal infection carried by small beetles. To stop the disease spreading, farmers were advised to cut down any infected trees. But the Government circular forgot to tell them to burn the wood! I do not know what the candidates might be for the most costly communication breakdown in the history of the human race — it is interesting to speculate — but I can tell you that our little beetles are still spreading merrily through England destroying elm trees as they go. And, how often have we met forms we could not fill in or regulations we could not properly understand? How often are service manuals unuseable? How often is our printed communication useless except to those who already know the answers? This sort of thing occurs too often and is too serious for us to accept complacently. We must do something about it.

For the past year my Research Assistant, Robert Waller, and I have spent many hundreds of hours investigating the ways in which complex printed communication could be improved. (We have reported some of the preliminary results of these enquiries in Macdonald-Ross, 1975; Macdonald-Ross & Waller, 1975a, b, 1976.) The work falls into two parts: that which concerns the research worker, and that which concerns the practical communicator — the editor, designer, author, or whatever.

The trouble with research is . . .

Researchers often talk as though they had a great deal of knowledge which was going to waste; they suggest that practical people outside universities are stubborn or pigheaded because they go their own sweet way and take no notice of what the researcher says. Now, it is very convenient for the researcher to put the word about, that research is very successful and is leading to deeper understanding and a changed world. But, what is the actual reality? Just how much is known and what may we expect of researchers? Let us look into the matter together with the help of some examples from the field of graphic communication.

I'll start with the research into typography. This is a well-established and apparently well-researched area; the key reviews are Pyke (1926), Paterson and Tinker (1940), Luckiesh and Moss (1942), Burt (1959), and Tinker (1963). We reviewed this work in detail and began to identify systematic defects which we documented and attempted to explain. In the first place, there are many experiments whose relevance is quite obscure. For example, there are many tests for the relative legibility of characters in a fount. Lists of "letters of low legibility" or "pairs of letters most likely to be confused" abound in the literature. But, no-one tells us what we should do with such results and I suspect this is proper caution, for it has long been known that occasional confusion of individual letters is not of prime importance in the reading process.

Much more important are the omissions. There is a sense in which legibility research has hardly started — so many important topics have been tackled. There are different kinds of omission. Sometimes previous research needs to be updated; for example, tests done on obsolete typefaces need to be redone on new typefaces. But a more important omission by far occurs when researchers overlook the advent of a completely new problem. For example, program learning brought with it questions and answers (prompts and responses); yet there is no work on how to set such items in relation to the rest of the text. Modern science textbooks rely on a close interdependence of text and diagram; yet, again, there is no research to guide the designer or the typographer. Errors of irrelevance and errors of omission are, of course, connected. The same set that predisposes researchers to choose problems that can be neatly handled in a laboratory, also prevents them going outside their circle to see the problems that occur in practice.

Many people have been and are still misled by the "classical experimental paradigm". For example, an experimenter wishes to test the effect of different line lengths and controls for size of type, interline spacing, and so on. Results are obtained and recorded in all the textbooks. But, altogether unnoticed, the researcher has used justified lines; that is, spacing between words varied so that the right-hand margin is kept aligned. Now, this interacts with the length of line: the shorter the line, the more the interword spaces vary; with lines five or six words long, some of the spacing effects can be quite bizarre. Very likely, such spacing effects

do bias line length tests in favour of longer lines. Yet, in all the classical line length tests,-justification was not "noticed" and hence, not controlled. There is actually now some evidence that ranged left (unjustified) settings are more legible for short lines and less able readers (Poulton & Gregory, 1970). This is small comfort, for without redoing dozens of earlier experiments, we cannot really be sure how to interpret all the earlier results.

We attribute such happenings not to the incompetence of researchers, but to their mistaken use of the classical experimental paradigm drawn from the physical sciences. Some researchers in recent years have tried to avoid this kind of problem by using the analysis of variance methods. They treat Anovar like a kind of trawl-net bringing in all the little fishes. The little ones get slung out and the bigger ones published. There is not much evidence that the advent of these modern statistical techniques has improved the quality of research; in some cases it has made it worse by putting a greater distance between the question the research worker poses and the answers he obtains.

When researchers go wrong, they go wrong in a big way. Empirical research is difficult, expensive, and time consuming. At any given time there will only be a few people conducting experiments into any particular problem area. If they misdirect their attention, many years may elapse before more fruitful work can be started.

*We have proposed a model for researchers that we think is more fruit*ful than the classical paradigm (see Figure 1). Our model requires them to make two preliminary readjustments. The first is to switch the object of research from discovering universal truths to improving specific decisions in particular practical situations. For example, the purpose of legibility research we see as to improve the quality of practical decisions by typographers and editors. There are excellent models for research of this kind: operational research and ergonomics are two examples. The second readjustment is to acknowledge the importance of tacit know-how. Any acquaintance with typographers and designers will show that their skills are backed by tacit knowledge which is the product of experience, rather than empirical tests or "book learning". Usually, researchers dismiss such personal knowledge or even deny its existence simply because it has not been developed by orthodox experimental methods. But actually, as Polanyi (1962) has shown, all scientific knowledge (indeed all human knowledge) has its roots in the "tacit dimension". So we should value the personal skills of typographers and designers and take them as the starting point for more fruitful legibility research.

Having made the preliminary readjustment, I turn now to the model. The model starts with the act of criticism. Criticism is a lynchpin for research, and it is one of the ways tacit know-how gets placed in the public domain. For example, when a designer writes a report for a cleint, he usually starts with a critique of existing solutions. Few problems are really novel and he can start by pointing out the strengths and

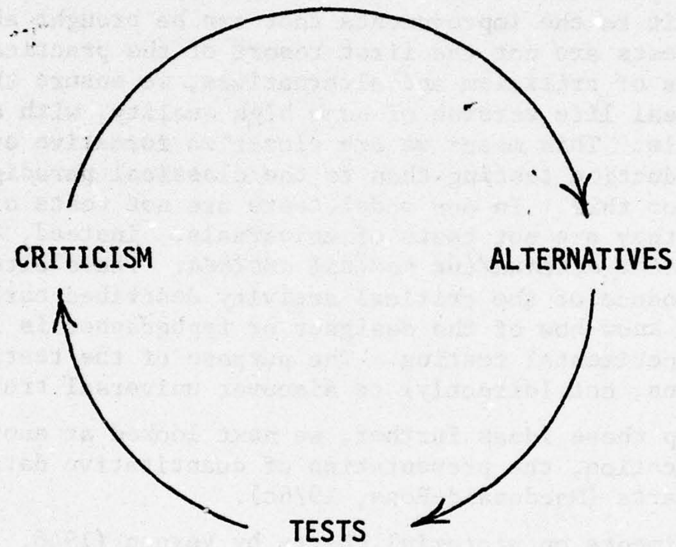


Figure 1. The Research Cycle.

weaknesses of previous designs, and thus, set up criteria for his own alternative solutions. Criticism is a basic intellectual skill. It is a primary method in the humanities and fine arts and, as Karl Popper has shown, plays an important role in the natural sciences. The habit of criticism or intellectual discernment is not only respectable, but actually essential for progress to be made.

The next part of the model is the creation of alternative solutions. These alternatives can be subjected to criticism and testing. The alternatives make use of the tacit knowledge which researchers have previously tended to disdain.

The third part of the model introduces empirical testing. Of course, there is a limit to the improvements that can be brought about without testing, but tests are not the first resort of the practical man. By the previous stages of criticism and alternatives, we ensure that the testing is done on a real life version of some high quality, with all the complexity that entails. This means we are closer to formative evaluation or industrial production testing than to the classical paradigm. And we make no apologies for this. In our model, tests are not tests of individual variables and they are not tests of universals. Instead, *the researcher tests one coherent alternative against another.* These alternatives are the refined produce of the critical activity described earlier. In this way, the tacit know-how of the designer or typographer is linked to the activity of experimental testing. The purpose of the testing is to make better decisions, not (directly) to discover universal truths.

To develop these ideas further, we next looked at another kind of visual communication, the presentation of quantitative data in tables, graphs, and charts (Macdonald-Ross, 1976c).

The experiments on pictorial charts by Vernon (1946, 1950, 1951, 1952a, 1952b, 1953, 1960) show some typical problems in experimental design. Vernon tested pictorial charts (which she called "Isotype") against tables and graphs, obtaining a no significant difference result. Her own conclusions show that she was prepared to take this result at its face value:

> No all-round advantage is gained by using pictorial charts rather than graphs. The charts appeal more to the less educated individual while the graphs may intimidate him; but it is unlikely that he will understand the charts sufficiently well to make it worth using them. (Vernon, 1952a.)
>
> It had been supposed that because these charts somewhat resembled ordinary pictures, therefore, anyone who studied them would be able to perceive and understand them, not as collections of little men, etc., but as data on the frequencies with which people performed the activities shown in the chart. It is true that, particularly when they first appeared, the novelty of these charts aroused people's curiosity, and therefore they possibly

> devoted more attention to them than they would have given to conventional diagrams or tables of figures. But, as the author found in her inquiries, a mere study of the charts and perception of the pictorial shapes they showed was no guarantee that their meaning would be understood. Such understanding necessitates a reasonable intelligence and special training in how to interpret them. (Vernon, 1962.)

However, there is a perfectly simple alternative explanation for Vernon's results, namely that the experiments were conducted with *incompetently designed graphic alternatives*. A test of a well-set table against a poor graph and a dreadful pictogram is invalid; one is not entitled to draw any general conclusions whatsoever. Moreover, when subjects fail, is this because only bright, well-trained people can understand pictograms? Or because it takes unusual intelligence to successfully unravel the visual confusion of a badly designed chart?

By good fortune, Marie Neurath (one of the original members of the Isotype Institute) is alive and well and living in London. Mme. Neurath has worked on pictorial charts for over half a century, so it seems fair to regard her as a "master performer". With her help we were able to construct a genuine Isotype chart from the original data used by Vernon. There can be no doubt that the Neurath chart is different from the Vernon chart and clearly superior as a piece of graphic communication (see Figures 2, 3, and 4). Of course, only detailed empirical studies can tell us the extent of this superiority — and in due course I hope to report the results of such studies.

To sum up, it is clear that the chief mistake Vernon made was to disrespect the art of chart construction. Her examples were simply not worth the bother of testing; they would have been greatly improved by the criticism and alternatives which our model prescribes. The results of her work, if meaningful at all, say to us only that badly constructed material does not communicate — and we hardly need researchers to tell us that! She, however, interprets her subjects' mistakes as due to their lack of intelligence and prior training. (Thus, is research so often contemptuous of ordinary people whilst being thoroughly uncritical itself.)

The passage of time has, alas, taught researchers very little and they continue daily to make the same kind of mistakes. As a result, important problems are not sufficiently researched, and when they are researched, the results are often not robust enough to stand up to application in the field. Yet I continue to believe that research is necessary and that these problems can be overcome.

From all these criticisms one can make a few exceptions. The field of readability research is noticeably more organised and more reliable than most others. Because it tackles problems which are derived from practice, and not from abstract theory, readability researchers do not usually have high status in the profession of psychology and may well not

NUMBER OF MALES AGED 14-64 IN GREAT BRITAIN

TABLE 1 — Thousands

Mid-year	Armed Forces (1)	Whole-time Civil Defence	Industrial groups — Group I (2)	Group II (3)	Group III (4)	Unemployed	Rest of male population (5)	Total male population aged 14–64
1939 ...	477	80	2,600	4,688	5,798	1,043	1,324	16,010
1941 ...	3,271	324	3,140	4,264	4,116	158	704	15,977
1942 ...	3,785	304	3,285	4,154	3,553	103	750	15,934
1943 ...	4,284	253	3,305	4,040	3,093	76	870	15,921
1944 ...	4,502	225	3,210	4,059	2,900	71	943	15,910

(1) These figures, and also the total column, exclude prisoners and missing.
(2) Munitions industries, *i.e.*, iron and steel, non-ferrous metals, shipbuilding, engineering, aircraft and vehicles, instruments, chemicals, explosives, etc.
(3) Agriculture, mining, National and Local Government, transport, shipping (including Merchant Navy), public utilities, food manufacture.
(4) Building, textiles, clothing, distribution, professional services, etc.
(5) Schoolboys, students, invalids (including war invalids), retired, etc.

13. Table 1 shows that while the numbers in the Armed Forces, whole-time Civil Defence and the munitions industries have increased, the number employed in other industries has been reduced. In 1944 the number of men engaged in the Group III industries was half of the corresponding pre-war total.

Figure 2. Source of Data for Vernon's Chart. (From the 1946 British Government White Paper, *Statistics Relating to the War Effort*).

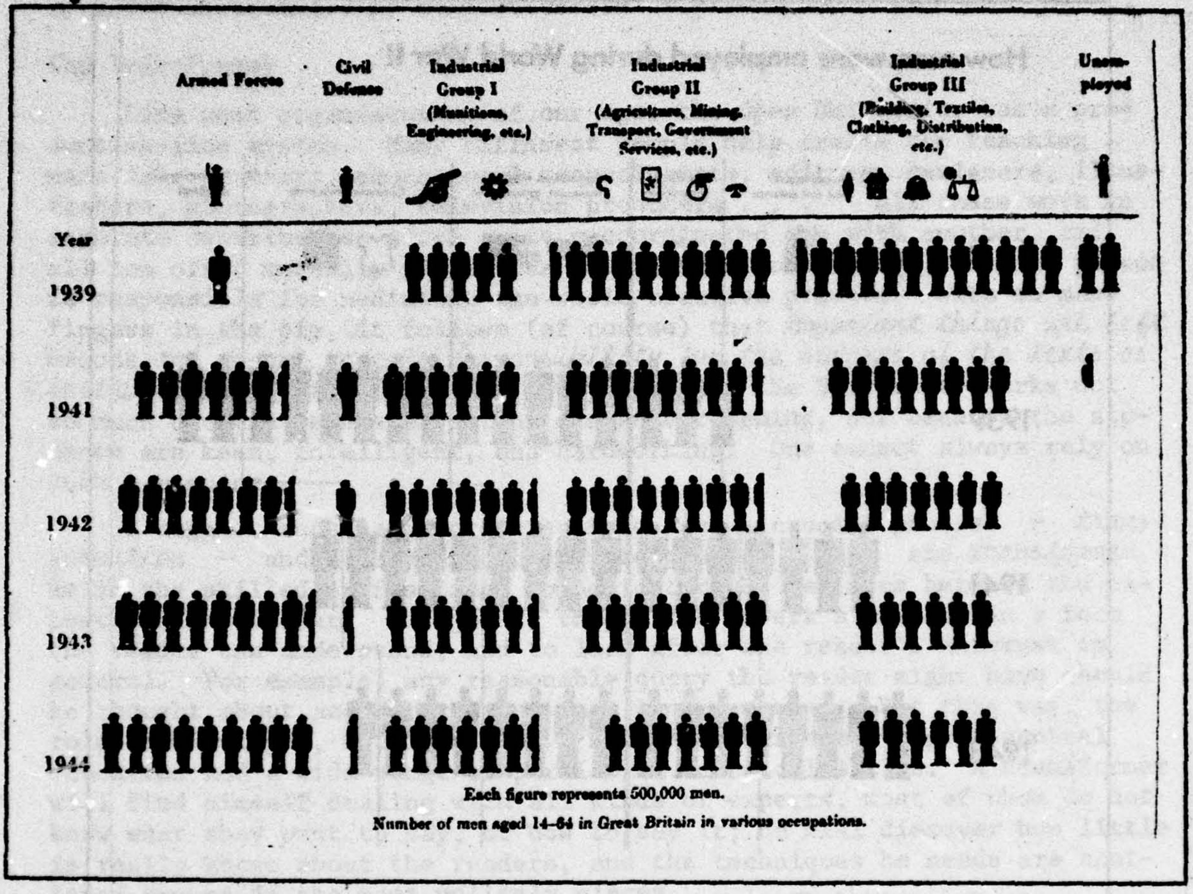


Figure 3. Vernon's Chart.

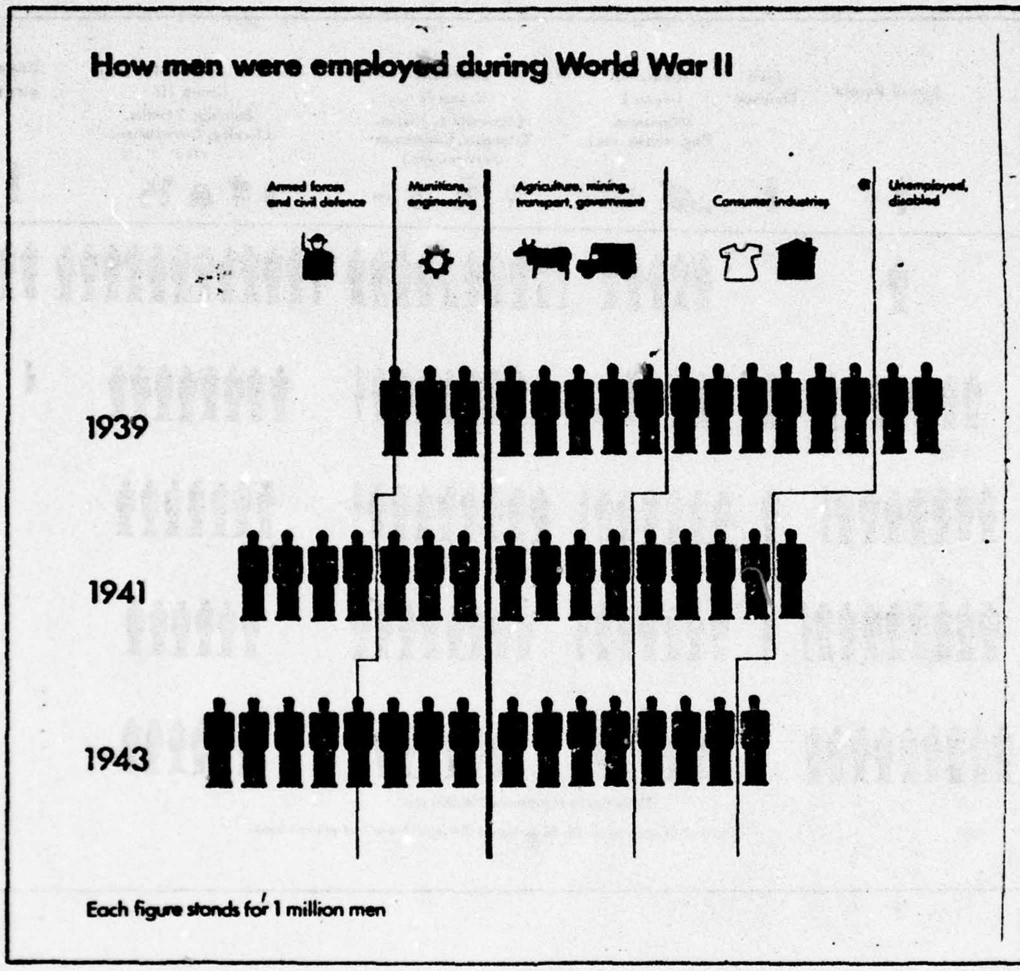


Figure 4. Neurath's Chart Based on Same Data.

get the credit they truly deserve. I do not say this simply because at this conference there are some of the very workers who have contributed to this field of research, but because it is my truly held conviction that other applied psychologists could learn how to redirect their efforts by close observation of this successful field.

The Transformer

Like most organisations of our age, the Open University has a production-line system. Many different people help create the teaching material: authors, educational technologists, editors, designers, illustrators, photographers, television producers . . . All these work in separate departments, often quite uncoordinated one with another, and all too often mutually suspicious rather than cooperative. No one person is responsible for mediating the whole creative process. With so many fingers in the pie, it follows (of course) that *important things are left undone and no-one accepts responsibility for the success of the texts as instruments of communication.* In my opinion, the University works not so much because we are expert at distance teaching, but because the students are keen, intelligent, and hardworking. One cannot always rely on such a response.

I suggest such large organisations lack a crucial process — *transformation* — and a person to carry the process out — the *transformer.* He is the skilled professional communicator who mediates between the expert and the reader. His job is to put the expert's message in a form the reader can understand, and to look after the reader's interest in general. For example, any reasonable query the reader might have should be thought about and catered for in a proper manner. Put this way, the role seems simple, but it is not. A transformer needs a good general education and a wide range of particular technical skills. A transformer will find himself dealing with all kinds of experts, most of whom do not know what they want to say, or how to say it; he will discover how little is really known about the readers, and the techniques he needs are scattered around in the most unlikely places.

I would like to have you think that the idea of the transformer is my own, freshly minted. However, it is not. It was first coined by Otto Neurath when he developed the system of graphic communication known as Isotype. Though I extend the idea beyond its original scope, Neurath's basic idea was correct, important, and . . . ignored for half a century! There are penalties for being ahead of your time.

The transformer starts with what to say, and then resolves how best to say it. Naturally, this distinction must not be overdone. What you want to say does partly determine how you say it and, in return, the content of a message is always altered to some extent by the way it is put over. Nevertheless, the distinction is a useful one.

First, one discusses the *content* of the message with the experts; later, one works out the exact *form* of the message with the help of illustrators, photographers, printers, and other technical staff. The skills needed and the problems faced differ at each stage. Finally, the cycle is complete when the transformer discovers what *effect* the message has on the reader.

The transformer is overseer of the whole process of communication — what is said, how it is said, and what its effect is. He works with colleagues whose skills are more specialised to make sure the message gets across and to reduce the change of communication breakdowns. He acts on the reader's behalf as best he can, sorting out the kind of issues a reader might raise if he were present in person.

<u>What To Say</u>

The transformer must be as closely involved with the content of the message as he is later with its presentation. He must understand the subject at hand so that his later judgements are *informed* judgements; but there is more to it than that. The transformer will usually need to help the expert get his ideas sorted out so that the subsequent act of communication has some chance of being successful.

The transformer is the partner of the subject-matter expert and not his slave. Our experience suggests that experts are not always as expert as they seem at first sight, and even published statistical data may be unreliable or misleading. Moreover, the expert's natural tendency is to think of his subject, rather than to think of the reader. For reasons of this kind, the transformer should not accept an author's instructions without critical thought. He must question and analyse until he can put the author's intentions in proper perspective. The idealistic young communicator may be surprised to find his source vague and confused instead of clear and authoritative, but the man is only human! He may not have organized his ideas yet; he may feel uncertain and anxious about issues that he has not sorted out but which cannot be ignored. The transformer helps as a sympathetic listener who gently refuses to go away until the confusion is sorted out. Any any rate, he must stick to it, for no-one can make a good communication out of muddled thinking.

The transformer sets about his task with the help of two insights: a good communication is *selected for a purpose* and has a *sound logical structure*. These two insights lead to the techniques which a transformer can use in his discussions with the source.

- Purpose and Objectives. All human communication depends on artful selection, since one can never say all that is known. The simplest basis for selection is a clear statement of purpose. The dialogue between expert and transformer may start with vague statements like: the reader should have some idea of this or should appreciate that. The transformer works to derive objectives which are as precise as the particular situation requires. For example, he may ask, *what will a successful outcome look like* ? What will the reader be able to *do* if he has interpreted the

message correctly? After clarifying what the expert wants to say and why he thinks it worth saying, one can discard irrelevant material and identify the prospective pay-off for a prospective reader.

- Tasks and Errors. Sometimes the work on the objectives must be supported by the kind of professional study often found in military or industrial training. If the reader is to perform a well-specified function or activity, then that activity can be analysed into its constituent tasks, and typical mistakes can be collected and analysed. Often, such a study starts with the "master performer" — an experienced person who exhibits all the skills and know-how necessary to meet the most demanding job standards. Such a study justifies the selection of objectives by *connecting the message with the reader's world.*

- Organising Principles. In education we make less use of the notion of tasks and errors and more use of organising principles. All subjects consist of facts, arguments, theories, problems, and procedures which (to a greater or lesser extent) are unified by central themes or organising principles. Sometimes these principles are quite grand, covering vast domains (the theory of evolution would be one such example). Sometimes the principles are more practical; for instance, computers can only carry out those operations which can be exactly specified in a predetermined code. Sometimes (as in the social and human sciences) one finds competing world-views which one must understand in order to interpret the opinions of the experts. Such world-views are to some extent *mutually exclusive* and a transformer must realise that they lie behind and largely determine what experts say and do. The principles help the expert, and can help the transformer, organise a map of the terrain.

- Facts. It is a fact that 8% of men (but only 0.4% of women) have defective colour vision. It is a *fact* because experts agree on what "defective colour vision" means: there are standard tests to identify it and the results of key investigations are not in dispute. On the other hand, if someone says "whites are more intelligent than blacks", the status of such a statement is much less clear. The statement is not known to be true; but it is not known to be untrue — it is in dispute. Therefore, the job of a transformer is to intercept opinions or interpretations masquerading as facts. He must identify the *status* of any key statements made by the expert.

- Truth and Validity. People differ in their ability to deploy arguments in an appropriate manner. This is too big a subject to be fully treated here, but so important that it must be mentioned. The transformer's dream is to see all arguments valid and all propositions as they claim to be (true, untrue, or uncertain as the case may be). Although this is the kind of quest which never quite ends, even a moderate amount of success does a good deal to ease the reader's problems.

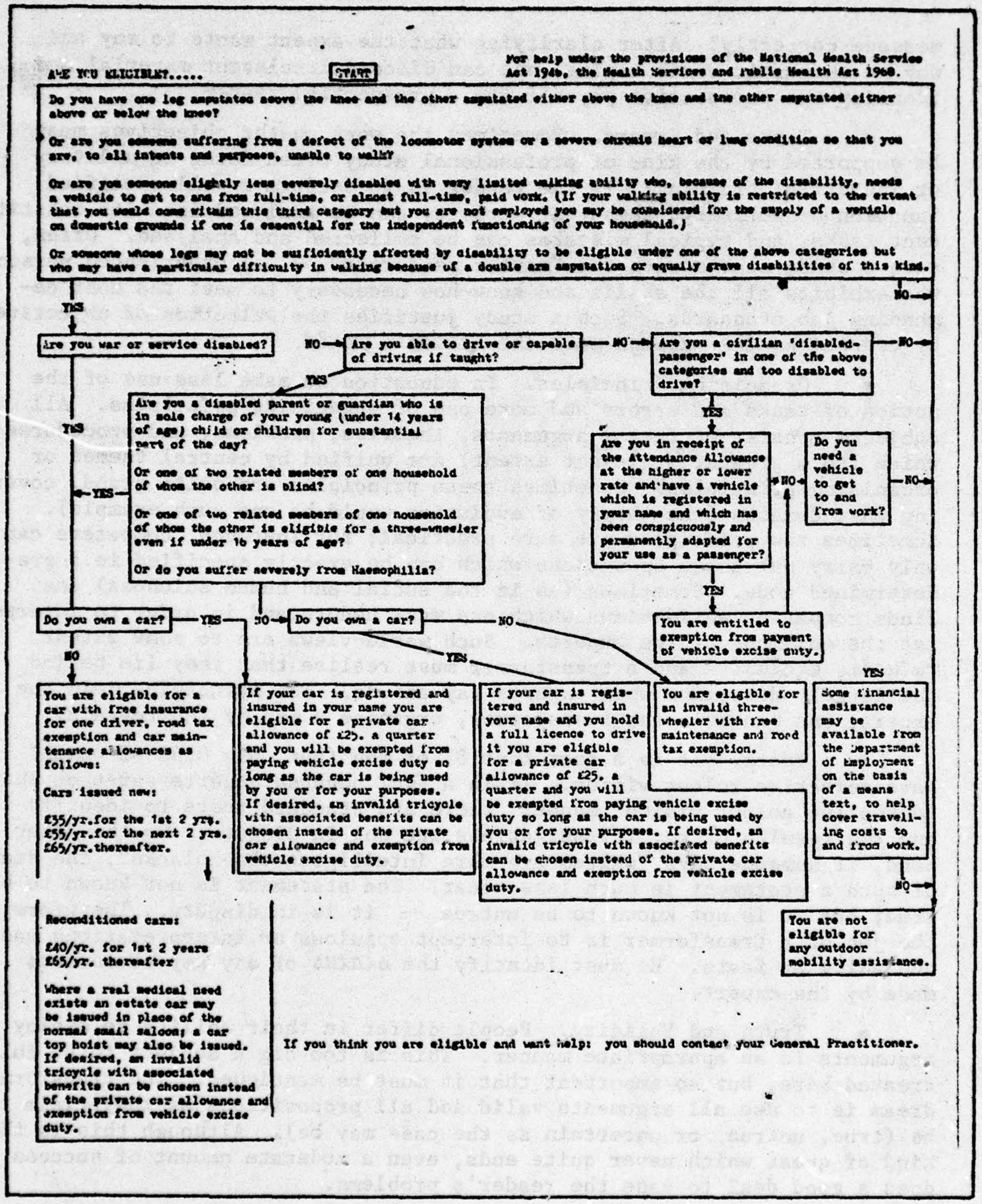


Figure 5. An Algorithm Before Transforming.

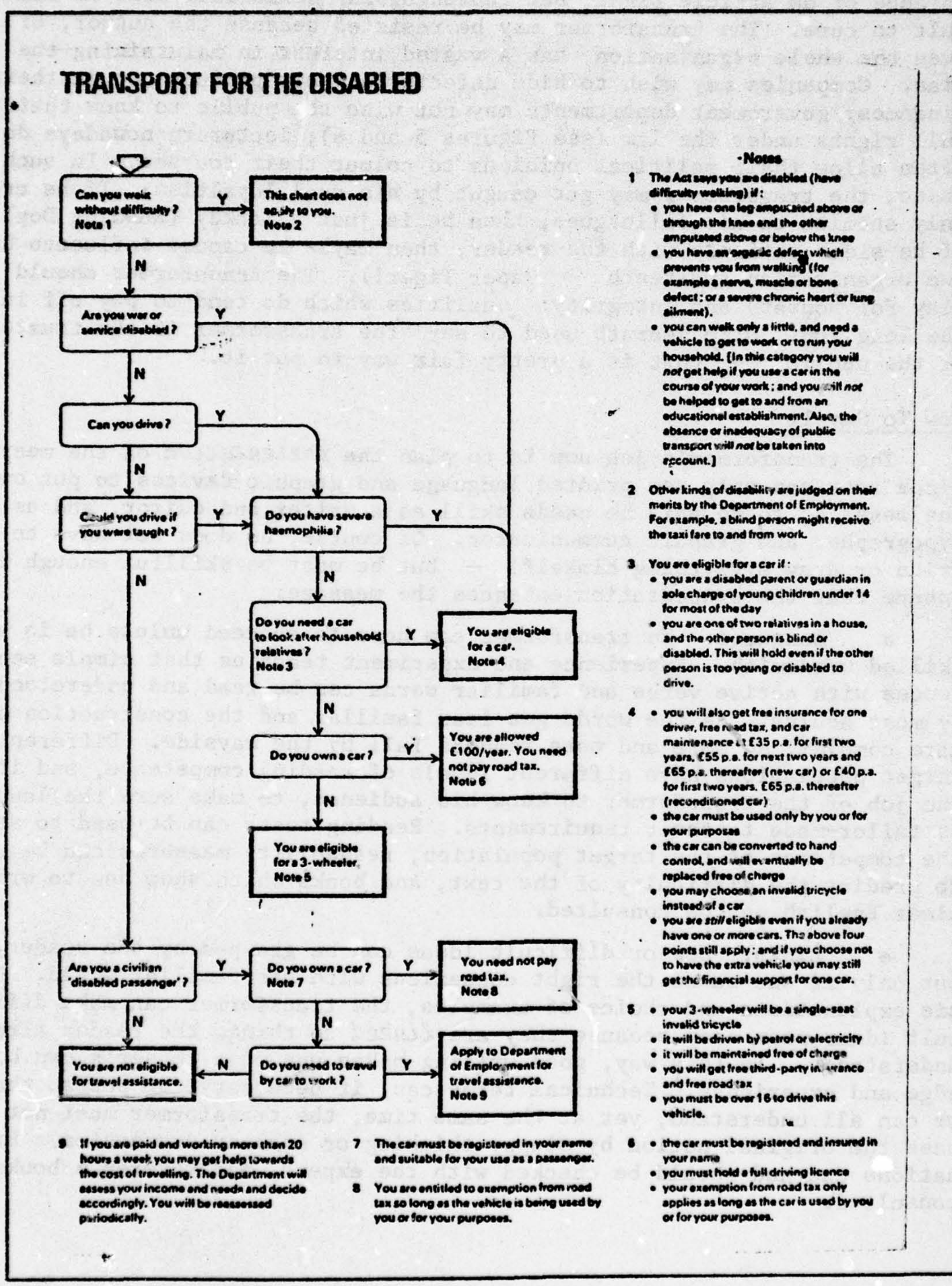


Figure 6. An Algorithm After Transforming.

• Bias. Most people welcome a second opinion to help get the balance of an article right, but thoroughgoing deliberate bias is difficult to cure. The transformer may be resisted because the author, or even the whole organisation, has a vested interest in maintaining the bias. Companies may wish to hide defects in their products or in their finances; government departments may not wish the public to know their full rights under the law (see Figures 5 and 6); lecturers nowadays do often allow their political opinions to colour their courses. In such cases, the transformer may get caught by his dual loyalties. If he cravenly submits to his colleagues, then he is just a lackey (Running Dog); if he sides entirely with the reader, then maybe he cannot influence his own organisation (no teeth — Paper Tiger!). The transformer should play for honesty and integrity: qualities which do tend to pay off in the long run. Otto Neurath used to say "the transformer is the trustee of the public" — that is a pretty fair way to put it.

How To Say It

The transformer's job now is to plan the *presentation* of the message. Since text can only use printed language and graphic devices to put over the message, to do this he needs skill as a writer and editor, and as a typographer and graphic communicator. Of course, he does not have to write or draw everything himself! — but he must be skillful enough to ensure that the presentation enhances the message.

• Language. No transformer can hope to succeed unless he is a skilled wordsmith. Experience and experiment teach us that simple sentences with active verbs and familiar words can be read and understood by most adults. As the words get less familiar and the construction gets more complex, so more and more readers fall by the wayside. Different target populations have different levels of reading competence, and it is the job of the transformer to know his audience, to make sure the language is tailor-made to their requirements. Reading tests can be used to assess the competence of the target population, readability measures can be used to predict the difficulty of the text, and books which show how to write clear English can be consulted.

• Links. New or difficult ideas can be grasped by the reader, but only if one makes the right connexions with his familiar world. By his explanations and choice of examples, the transformer can make difficult ideas seem easy because they are *linked* to things the reader already understands. In this way, good writing makes use of a reader's own knowledge and experience. Technical terms can, if necessary, be put so that we can all understand, yet at the same time, the transformer must not debase the original notion by sloppy thinking or unclear expression. Explanations can and should be checked with the expert, and reference books consulted.

• Typography and Graphics. To attract and keep interest and to ensure legibility are the classic aims of the designer and the typographer. These aims are still important, but the transformer must have a much broader vision than that. Often, the graphic elements carry a crucial part of the message, so one cannot regard design as just external decoration applied to an existing message. Rather, we go along with the ideals of the New Typographers: design is an integral part of the communication process. Good typography helps the reader plan his reading strategy, tells him where he is and helps him to find his way about; good graphic design allows one to say in words *and* illustrations what could not be said in either form alone.

• Action. If one wants the reader to *do* something after reading the text, then these points are crucial. We assume that *what* the reader should do is already decided (by task analysis or whatever); now we want to know if the text enables him to perform. The three key questions are — will he understand what we are asking him to do? Can he actually do it? (is it physically and psychologically possible?). Will he know when he has completed the task successfully? The simple Highway Code rule "RED means *stop*" meets all three criteria simultaneously in three words. But, what about "at 70 mph shortest stopping distance is 70 ft thinking distance + 245 ft braking distance = overall stopping distance 315 ft"? This is useless information. In the first place, the accuracy is spurious, as the Code admits (wet and slippery roads, different vehicles, poor brakes and tyres, tired drivers . . .). But, in any event, how on earth are drivers meant to translate those numbers into appropriate action on the road? The answer is, they can't and they don't! Good drivers may learn to judge stopping distance reasonably well, but no thanks to the Highway Code for that.

• Organisers and Signposts. The old saying goes: "First I tells them what I'm going to tell them; then I tells them; then I tells them what I've told them!" People do need to know where they are, where they are going, and what the prospective pay-off is. As we proceed, they need to know what the *status of the message* is. Are we giving them main or subsidiary information? Are we asking a question or giving an instruction? The form of the message must show its status and function. Texts are full of devices which help the reader find his way around; typical examples are objectives, prefaces, introductions, contents lists, headings, questions, instructions, numbering systems, glossaries, and indices. If the communication aims at specific goals, then the organisers are strongly directive (objectives, instructions); if the reader is allowed more freedom, then the permission-giving organisers becomes more important (contents, headings). Thus, the choice and emphasis depends on the purpose of the communication and the situation of the reader. Organisers are not isolated bits and pieces; they must fit together. The contents page, glossary, index, numbering system, headings, diagrams, and main text are mutually interdependent; what one does in one place affects all the others. The design and layout of a two-page spread can become a little cameo of transformation; all kinds of informed guesses are made about reading behavior so that the reader can best access the information he needs.

• Feedback. Since we cannot be there to answer questions, the reader may often wonder whether he has understood us properly. So we must allow him to check his understanding by giving him feedback. We must tell him how to recognise when he has executed an instruction or answered a question correctly or understood a key idea. The feedback should debug typical errors (take care of the most frequent kinds of mistake) and should be placed immediately after the stimulus (question, instruction, or whatever). It is worth mentioning that feedback can be quite subtle; for example, repeating the same idea in different form gives the reader a second map-bearing which checks his first interpretation — "Of making many books there is no end; and much study is weariness of the flesh" (Ecclesiastes 12, v 12).

• Testing. The transformer *cannot* get everything right; he is lucky to get things half-right at first. Yet the final version may need to be very good indeed, if it is to succeed. How do we face this problem? We say, there is a time for guessing and *a time for testing*. The testing should be done during the transforming process on a rough but complete draft version. The object is (obviously) to gauge whether the purpose of the communication is being achieved, and to identify and collect sources of misunderstanding. Any text which is important enough to have money spent on it is important enough to be tested, and the transforming process must allow time for the lessons learnt to be put into effect.

The Later Stages

Earlier we saw how the transformer worked with the expert source; now we see him working with two quite different kinds of persons — the skilled craftsman (or technician) and the research worker. We deal here in less detail than earlier, though in the long run the transformer's relations with craftsmen and researchers will be just as important as his relations with sources.

• Production. The transformer depends upon illustrators, photographers, compositors, film editors, and other skilled persons to help him realise his vision. These people can make a real contribution. At their best, they can teach the transformer a good deal about the possibilities and limits of their arts, and about the costs and implications of various production methods. In some cases (scientific photography, for example) a skilled person can help the growth of the subject-matter as well as its subsequent presentation. The transformer should deal with these people in person, not through any kind of intermediary. It is worth mentioning that without the transformer, the final product often reflects the production process rather than the needs of the reader. How often have we seen the fussy house-style applied unintelligently or the diagram which has lost its meaning because the illustrator did not understand it! Often a free-hand sketch by someone who understands works better than that same sketch redrawn "professionally" by an illustrator. Underground papers use modern print methods to *bring the communicator closer to the reader*. We should think long and hard about such ideas.

• Evaluation. "Discovering the effects" is the last part of the transforming cycle. This means going out into the field and looking at *all* the consequences of the communication that can be found. The really gross errors should be picked up earlier in developmental testing, but many subtle effects only show up later in the field. The information gained can be used for subsequent editions, but is in any event useful, since similar problems and situations are just bound to occur. Transformation is an active process that learns, not a passive dispensing of recipes. Evaluation can be a complex technical matter. Unless the transformer himself has research training, he may well need the services of a professional researcher to alert him to ways of collecting data. The budget for evaluation depends on the cost and importance of the communication and the prospects of similar publications in the future; we do suggest that there should always be such an item in the budget.

• Recurrent Problems. As time goes on, the transformer will notice certain problems coming up time after time. Sometimes they cannot be solved simply or intuitively; they need detailed work in depth for a long time. This is the job of the basic researcher. The relationship between these two colleagues is often uneasy, with the transformer wanting solutions yesterday and the researcher limited by what it is *possible* to do, given present knowledge and resources. A similar case is the relation between general practitioners and medical researchers. The GP has patients *now* in his surgery, but the researcher knows that some questions just don't have easy answers. For all that, the art of medicine advances by this mixture of cooperation and conflict, and so should the art of communication.

PART II

REPORTS ON READING AND READABILITY RESEARCH AND DEVELOPMENT IN THE ARMED SERVICES

The views expressed in these papers are those of the authors and do not necessarily reflect the views of their respective military services.

Preceding page blank

HISTORICAL PERSPECTIVE
ON RESEARCH AND DEVELOPMENT ON LITERACY AND
TECHNICAL WRITING IN THE ARMED SERVICES

Preceding page blank

INTRODUCTION

The current research and development programs in the areas of reading, readability, and useability in the Armed Services have built upon earlier efforts in these areas. This section looks back on that earlier R&D to help provide a perspective on the current R&D.

In the first paper of the section, Dexter Fletcher highlights developmental efforts in his historical survey of literacy training in the Navy. Looking at the years from the 1700s to the outbreak of the Second World War, he discusses the early literacy training, which was prompted in part by religious reform and in part by the naval apprentice system. He then presents a more extensive discussion of the three literacy training programs of the period following 1943: the Special Training Program, which was prompted by World War II; the Recruit Preparatory Training, which was prompted by the Korean Conflict; and Academic Remedial Training, which was prompted by Vietnam and Project 100,000.

Fletcher emphasizes the need for systematic literacy research and development, rather than periodic development in times of emergency, which has been the case in the past. He also points to the increasing concern for job functional literacy to enable Navy personnel to perform their jobs.

The emphasis switches from development to research as Jack Hiller surveys literacy research in the Army. He uses a systems approach to uncover pertinent research, and the lack of it, in the areas of reading training, technical writing, and useability of job and training materials. His historical review points up the fact that research has lagged far behind development in reading and readability R&D in the armed services.

The history of literacy research and development in the Air Force of necessity begins later than that of the other services. In presenting that history, Ron Burkett reminds us that the problem has always been a two-sided one of both readability of the materials and reading ability of the personnel. He discusses both sides of this problem. On the one side, he presents the efforts to determine the readability of Air Force publications, both job and training materials, and the attempts to make those materials more readable. On the other side, he looks at studies which attempt to determine the reading abilities of Air Force personnel and those which document the match or mismatch between the reading ability of personnel and the readability of the materials. The documented mismatch, or literacy gap, in 10 problem career fields is approximately two reading grade levels.

Burkett also looks at some of the ways in which the Air Force has dealt with the problem. He reports on personnel studies, which have been used to support a policy of minimizing the literacy problem through high selective standards; evaluation studies of the literacy training programs, which attempted to close the literacy gap by raising the reading abilities of the personnel; and developmental efforts to produce the job performance aids, which attempt to lower the reading difficulty and increase the useability of job materials.

HISTORICAL PERSPECTIVE ON LITERACY TRAINING IN THE NAVY

J.D. Fletcher
Navy Personnel Research and Development Center

History seems to be not so much a record of what has happened as a record of what we remember, and any historical perspective is most probably a perspective on perspectives. This is certainly true of the current historical perspective on literacy training in the Navy. Because of the subjective nature of the data that support it, this perspective is divided into two parts: a very brief part dealing with all of Navy history prior to 1943 and a more extensive part beginning in 1943 and ending with the present. Further, very little information is presented on readability and technical writing. Historical information on these topics doubtless exists, but it is buried in the instructions and official correspondence of the Navy Department.

A comprehensive, thoroughly satisfactory definition of literacy training was not attempted in this report. A rough description of literacy training as an attempt to bring personnel who read below the 4.0 grade level up to that level appears to serve fairly well the needs and scope of this report. In the history of Navy training, this description emphasizes training for enlisted men who are in their first tour of duty.

A. Literacy Training Before 1943

Enlisted personnel are, at best, neglected in military histories. Battles, tactics, and technologies tend to be recorded in some detail, but the ability and character of the enlisted force that are basic to any military service are often obscured by the statistics of manpower supply, loss, and demand. If it is reasonable to assume that enlisted personnel acquire a greater significance as the technological demand of their duties increase, then their neglect in histories of the United States Navy is particularly regrettable. It is difficult to read any account of enlisted life without bringing away an impression that Navy duties have always demanded substantial technological capabilities. The introduction of steam and electrical systems aboard ships obviously increased the technological demands of enlisted Navy jobs as Cummings (1929), Harrod (1973), Potter (1913), and others have indicated, but discussions such as those of Luce (1890) and Niblack (1891) of training problems that existed in the earlier, sailing fleet indicate that substantially more than disciplined responses to orders and knowledge of nautical terminology were relevant objectives in transforming "landsmen" into seamen. Most of this training was accomplished on the job following the apprentice, journeyman, master craftsman model. It required listening but not reading skills, and it continued

throughout a sailor's career. Despite the predominance of this mode of training there were significant efforts as early as the first half of the 19th century to provide instruction in the basic skills of reading, writing, and arithmetic to Navy personnel. There was no land-based training establishment at this time and basic skills training took place entirely aboard ship, at dockside, and while the ship was underway. There appear to be two primary reasons for the early literacy training: the religious reform movements of the 1700s and the assignment of teen-age boys to the fleet.

1. Religious Reform

In the late seventeenth century and during much of the eighteenth century, a major transformation in Western civilization took place under the auspices of the Enlightenment. Leaders of this movement placed great faith in human reason backed by the findings of science and submitted to question all authority and absolute standards. As a result, a sympathetic and humanitarian outlook on the condition of all men was stimulated, especially by religious manifestations of this movement. Protestantism particularly fostered an enlightened self-interest on the part of its followers by emphasizing good works and the notion that every man is his brother's keeper. Heaven would be won and past wrongs atoned for, if each man would firmly embrace religion and help his fellow man, especially his less fortunate fellow man, to do the same. Some members of the evangelical movement, which grew naturally from the Enlightenment, chose to focus their interest on the sailors of the naval and merchant service. Revivals were held, Sunday schools were conducted, and tracts were distributed. With the distribution of the tracts came the realization that the word of God was inaccessible to many sailors because they could not read. Moreover, sailors' moral welfare appeared endangered because their lack of arithmetic skills made them easy prey for the peddlers of prurience who were naturally drawn to the full pocketbooks and poorly educated minds of newly disembarked sailors. So it was that the first to minister to the need for essential skills training among Navy enlisted men were the early Navy chaplains (Langley, 1967).

The regulation of 1802 described the Chaplain's duties as the following:

1. He is to read prayers at stated periods; perform all funeral ceremonies over such persons as may die in the service, in the vessel to which he belongs; or, if directed by the commanding officer, over any person that may die in any other public vessel.

2. He shall perform the duty of a school-master; and to that end he shall instruct the midshipmen and volunteers, in writing, arithmetic, and navigation, and in whatsoever may contribute to render them proficients. He is likewise to teach the other youths of the ship, according to such orders as he shall receive from

the captain. He is to be diligent in his office and such as are idle must be represented to the captain, who shall take due notice thereof. (Burr, 1939, p. 111.)

In addition to Chaplains, there were schoolmasters and teachers in the Navy as evidenced by the law of March 1799 assigning them three-twentieths of prize money, which was approximately the amount assigned to warrant officers. In general, the schoolmasters made a poor showing. Allegations of sloth and drunkeness on their part appear well founded. Schoolmasters' duties varied with commanders, but they were generally charged with instructing the boys, apprentices, and midshipmen assigned to ships. The employment of schoolmasters was never extensive in the Navy, and it gradually died out. However, the history of Navy chaplains is far more honorable than that of the schoolmasters, and the Chaplain Corps maintained its interest in essential skills instruction for enlisted men throughout the history of the Navy (Drury, 1949). There was, of course, little systematic method in this instruction, and its success most probably keyed on the moral rather than cognitive well-being of its students.

2. Apprentices

From the beginning of the Navy there were boys assigned to ships. Attempts to standardize their employment and treatment were embodied in the apprentice systems set up in 1837, 1855, 1863, and 1875. Of all the training ventures before the Spanish-American War, only the naval apprentice program provided a prototype for the modern Navy (Harrod, 1973). Despite this fact, the apprentice system suffered unpromising beginnings. The 1837 effort failed because apprentices' hopes of obtaining commissions were not fulfilled and because of an alleged mutiny in 1842 on the training ship *Somers*, in which a son of the Secretary of War was hanged, probably wrongfully (Langley, 1967). The 1855 program was interrupted by the outbreak of the Civil War. The 1863 effort was unsuccessful because apprentices were again disappointed in their hopes of obtaining commissions. However, the apprentice system continued to be revived because of very serious problems the Navy was experiencing in manning American ships with native-born seamen (e.g., Luce, 1974). In 1875 the Navy began enlisting boys 14-18 years of age to serve as apprentices until age 21. The Secretary of the Navy emphasized that it was not the object of the system to prepare boys for commissioning; the apprentices were to be trained in all the duties of sailors on a man-of-war, and, significantly, they were to receive an elementary English education (Harrod, 1973). Despite numerous problems, including the neglect of academic training, the system, with its commitment to literacy training, survived until it was replaced in 1904 by an officially established apprentice seaman rating with a minimum age requirement of seventeen. With this system evolved the prototype for today's recruit training.

Recruit training was essentially motivated by a need for standardization in training landsmen and apprentices for shipboard duty. A natural means for standardizing training that was employed very quickly after the need for it was recognized was the publication of drill books. Fullam's *Recruit's Handy Book* and McLean's *Bluejackets' Manual* both appeared in 1902; Fullam's book lasted until the 1920s and McLean's manual is currently in its nineteenth edition. By General Order 114 of November 1902, the Navy Department required all recruits to know the contents of the *Recruit's Handy Book* and issued a copy to each. Fullam also published in 1902 the *Petty Officer's Drill Book.*

The appearance of these handbooks signaled the fact that oral instruction was no longer sufficient for an enlarged Navy (cf., Harrod, 1973). Presumably, this fact was already apparent. Some of the technical equipment aboard Navy ships at the beginning of the century was fairly sophisticated, and it seems likely that this equipment was accompanied by essential manufacturer's literature on its operation, maintenance, and repair. However, there appears to be little record of the nature, preparation, and supply of this technical documentation. In any case, the appearance of the drill books signified an official expectation that seamen would be able to read, and literacy effectively became a requisite Navy skill.

Despite a national requirement for literacy, about 11.3 percent of the United States population in 1900 could be classified as illiterate (Harman, 1970). There were no systematic attempts to screen Navy recruits until 1925 (Harrod, 1973), and illiteracy among Navy personnel became increasingly serious. Although there is no direct reference to this effect, the systematic screening that began in 1925 included written tests, successful performance on which required literacy. It seems likely, then, that the incidence of illiteracy in the Navy was substantially reduced by the use of these tests.

Aside from the work of the Chaplains, the only formal literacy training that appears to have been supported by the Navy in the period 1900-1943 was in Portsmouth Prison where some effort was made to rehabilitate men who were classified and jailed as deserters because of their inability to read furlough orders (Potter, 1918). These men would go on leave and return on what they thought was the appropriate date only to find themselves scheduled for court-martial.

Despite the screening process begun in 1925 and official policies for their exclusion, illiterates continued to appear in Navy billets throughout the entire period (Special Training Program, 1951). Other than the rehabilitation program at Portsmouth Prison, no official efforts were made to help these men achieve literacy. There was, and is, considerable unskilled labor required in many fleet billets, and these men were typically assigned to do this labor. As might be expected, rates of promotion and re-enlistment among these men were not high. During World War II, manpower supply began to run substantially short of demand and in June 1943 the Navy reluctantly agreed to accept its fair share of illiterates under Selective Service.

B. Literacy Training in 1943 and After

1. The Special Training Program

Although prior to June 1943, there was no official recognition of the fact that room would have to be made for enlisted personnel who could neither read nor write, a great many men in this category were already serving in the Navy. Immediately after Pearl Harbor, recruiting stations were released from the obligation to administer the General Classification Test (GCT) to applicants and since this regulation remained in abeyance until the start of 1945, many illiterates found their way into the Navy through traditional recruiting channels (Special Training Program, 1951). In August 1945, the Director of Training indicated in the following summary some of the problems the Navy and illiterates experienced with each other.

(1) At times the period allowed for recruit training was contracted by the demands of the service to four or five weeks. Under these circumstances, the trainee was obliged to acquire a large part of his instruction through reading. It was found that it took approximately four times as long to train an illiterate to perform an average Navy job as it did one who could read.

(2) The establishment of a training program which did not depend on the use of printed matter would have been both difficult and expensive. Experience showed that it was simpler and more economical to teach men to read than to devise materials which did not require this knowledge.

(3) The establishment of a smooth administrative routine was grossly complicated by the presence of non-readers. A system for the rapid handling of records was a virtual impossibility where men could not fill out information blanks, pay receipts, proficiency slips, allotment cards, et cetera.

(4) Sufficient education to read safety precautions was essential for men working with machinery, high explosives, and heavy cargo. Serious accidents were traced directly to men's inability to read warnings and study safety instructions.

(5) A social barrier of serious implications was found to exist between literate and illiterate personnel.

(6) The administrative dualism resulting from putting literates and illiterates together caused confusion. Literates tended to resent the long oral directions which they had to listen to for the sake of the illiterates in their number.

(7) A very large number of minor disciplinary problems were the direct outgrowth of misunderstandings caused by inability to read station orders, watch bills, leave and liberty regulations, and safety instructions.

> (8) An inability to read and write letters constituted among illiterates a serious morale problem and consequent obstacle to satisfactory adjustment to naval life. It became increasingly evident that a knowledge of reading and writing helped to overcome a feeling of inferiority and tended to develop initiative, aggressiveness, and more willing acceptance to the conditions of military life.
>
> (Special Training Program, 1951, p. 2-3.)

Despite this analysis, which was written after the fact, there is substantial evidence that authorities were reluctant to accept responsibility for setting up a special literacy program long after the need for such a program became obvious. However, as the situation gradually worsened, the Navy stumbled into a situation that required recognition of the need for a special literacy program. On 30 September 1943, the Naval Training Section in Norfolk, Virginia, reported that illiterates were being received from boot camps in increasing numbers. Since all the facilities at this section were needed for the destroyer and destroyer escort training programs, permission was requested to transfer these personnel to the Naval Training Section at Bainbridge, Maryland, for further instruction. Evidently, the permission was granted without due consideration of the practical consequences of this decision. The result was that Bainbridge found itself deluged with illiterates and appealed to the Bureau of Naval Personnel for help. The immediate response of the Bureau was to cancel the permission that had previously been granted to Norfolk. However, as a result of these events, the need for special literacy training was brought to command attention (Special Training Program, 1951).

The presence of a growing body of illiterates in the Navy created a training problem that it was totally unprepared to face. Luckily, as a thorough review of World War II literacy training programs by Fattu, Mech, and Standlee (1953) shows, the foundation for solving this problem had been laid much earlier by the work of the Civilian Conservation Corps (CCC).

The CCC was established in March 1933 primarily as a means for providing productive employment for young men during the depression. Although it was administered by the Army throughout its nine-year existence, much use was made of professional educators in developing and guiding the educational phases of the program. The CCC education program was secondary to the work program, and the literacy program was only a small part of the education program. It was created to serve the estimated six percent of CCC enrollees who were functionally illiterate in the camp life situation (Couch, 1944). Nevertheless, the literacy training program did exist as an identifiable entity and it served as a foundation for the development of similar programs, first by the Army in 1941 and later by the Navy in 1944 (Fattu et al., 1953).

The Army program was discussed by Heath (1946), who pointed out that there were three distinct groups of illiterates targeted by the Army program: English-speaking illiterates, non-English speaking literates, and Oriental literates. Training was tailored to the special needs of each of these groups. The teaching program for the English-speaking illiterates passed through five distinct phases. These phases were sequenced and a student had to master each phase before proceeding to more advanced ones. Heath describes the phases as the following:

1. Consonants and Key Words. Sounds of the consonants were taught in this phase by associating 21 key words with English consonants. Phonemes such as /š/, /č/, /ž/, /ǯ/, /θ/, and /ŋ/ were not taught.

2. Monosyllables and Vowels. Nonsense syllables such as ter, ker, and nub were initially taught in this phase in an effort to train students in the relationship between orthography and sound.

3. Introduction of Polysyllables. Students were taught to analyze polysyllables into monosyllables.

4. Sentence Structure and Word Recognition from Context. Further instruction in word attack was presented and instruction in constructing sentences and learning (recognizing) words from context was added in this phase.

5. Composition and Expression. Students were taught to read and prepare military messages and personal letters in this phase. They were also taught the use of such basic resources as dictionaries and telephone directories.

Discussion of the Army's literacy program is relevant because when the Navy finally faced the need to produce a literacy training program, it turned in September 1943 to the Army for guidance, and the workbook material developed for the Navy was based on the five phases discussed above.

On 22 December 1943, the imminent appearance of a special literacy program for recruits was announced by the Navy. On 23 December 1943, the Naval Training Section at Great Lakes, Illinois, was directed to prepare for the arrival on 3 January 1944 of an initial draft of 420 White illiterates. In March 1944, two Navy programs for literacy were established: one at Camp Peary, Virginia, for Whites, and one at Great Lakes for Blacks. Plans for the Camp Peary program called for a weekly input of up to 500 trainees with a total capacity of 6,000. In fact, the total enrollment at Camp Peary quickly grew to 10,000 in April 1944 (Special Training Program, 1951). Notably, both these programs were set up as an integral part of recruit training; no programs were established for illiterates who were already in the Navy. At no time did the Bureau of Personnel formally accept responsibility for training illiterates who were above the recruit level, although informal support was given to Commanders who wished to aid illiterates under their command with the dissemination of literacy training materials throughout the fleet.

The curriculum that was initially devised for the literacy programs in early 1944 provided for only 133 hours of instruction in reading and writing and 73 hours of instruction in arithmetic, for a total of 206 hours of instruction out of the 576 hours originally called for. However, by January 1945, a considerably expanded and improved program had evolved. Four basic assumptions not previously annunciated determined the form of the Special Training Program, as it came to be called. First, it was a training program for adults. Although unable to read and write, the trainees came to the Navy having command of a well-established oral vocabulary, together with a fund of experience that put them beyond the appeal of grade school readers. Second, the trainees were the products of a wide variety of environments, so that the only interests they could be expected to have in common would grow out of their life in the service. Third, the limitations on time allowed for acquiring literacy made it mandatory that the program be rigidly functional in nature. A nominal proficiency grade of 5.0 in reading and writing was established as the teaching goal, but the purpose behind the program was simply to qualify men to read, write, and figure sufficiently well to perform all essential Navy duties. Thus, it was expected that graduates of the course would be able to read watch lists and safety precautions, and would be able to fill in beneficiary slips and small store chits. On the other hand, there would be no attempt to indoctrinate trainees formally in such refinements as capitalization, spelling, and punctuation. Fourth, it was assumed that the vast majority of instructors would be inexperienced in the type of teaching expected of them. They could be expected to rely unduly on lecture and blackboard methods and to encourage parrot-like memorization of lists of words without developing the skills of analysis and synthesis that are essential to literacy (Special Training Program, 1951). The CCC literacy materials were called the *Camp Life* series, the Army materials were called the *Army Life* series, so naturally the workbooks, basic readers, supplementary readers, tests, and teacher's manual developed for the Navy were called the *Navy Life* series. Private Pete was replaced by Seaman Sam.

Development of the program materials was described by Ross:

> In writing this program, some radical departures from the conventional were taken, because the situation and the nature of the students and instructors demanded them. For example, a "reader" in the hands of an untrained instructor at the outset of the program would result in a static classroom situation in which the students "read" orally in rotation, with prompting, until the page has been virtually memorized. In the Navy Life series, therefore, the first book is not a "reader", but a workbook-type text which forces the instructor down from the platform among the students. The first of the readers is not introduced until considerable reading ability is developed through chart, blackboard, and workbook reading experiences. When it is introduced, no new skills or words are required for some time, and the student can read it easily for meaning. As a result, the student is literate so far as the readers are concerned from his very first experience with them.

> Comic books are exceeded in popularity by no other reading material. So, later in the program, when they can be handled easily, comic books are introduced for rapid, supervised classroom supplementary reading. These are regular commercial comic books, carefully selected in advance of publication, and then rewritten with a core vocabulary basic to the Navy Life program.
>
> (Ross c.s., 1946, p. 204.)

Notably, the cornerstone of this program was not the basal readers, which was probably the case in civilian initial reading instruction, but the program's two workbooks functionally entitled *Navy Life Book I* and *Navy Life Book II*. Both books were constructed on a basic vocabulary that derived first from words shown by research (what specific research is not in the record) to be an essential minimum for literacy, and then, to as great an extent as possible, from naval terminology. Fans of Leonard Bloomfield will be interested in Figure 1, which is the first "story" that occurs in *Navy Life Book I*. The emphasis on the single spelling pattern "at" embedded in hat, mat, and sat will seem very familiar to those acclimated to Bloomfield's "Nan can fan Dan" (cf. Bloomfield & Barnhart, 1961). Extensive reliance was placed on illustrations to indicate word meaning, but other devices such as a visual acuity test and illustrations of phonetic similarity were also used. It was assumed that by teaching reading and writing simultaneously, growing proficiency in either would increase proficiency in the other. The material in *Navy Life Book I*, which contained 400 illustrations, took as its common denominator barracks life, which all trainees might be expected to have in common. In the second volume, the emphasis was gradually shifted from phonetic elements to syllables and from illustrations to context as a means of furnishing clues to word meaning. The rigid control over the material that characterized *Navy Life Book I* was gradually relaxed in *Book II*, and was finally loosened so that any student capable of handling it would be able to conduct independent reading outside the classroom. *Navy Life Book II* was also prepared with the intention that it should be a useful adjunct to regular recruit training, since it contained a good deal of incidental information on such subjects as semaphore, firefighting, elementary navigation, naval customs, guns, ship types, seamanship, naval terminology, and personal hygiene.

As it evolved, the curriculum was developed on a flexible scale, operating between the limits of twelve and twenty weeks. In other words, trainees received periodic tests starting with the eleventh week, and could be graduated any time thereafter by showing that they had achieved the required level of literacy. The minimum overall time prescribed for the course was 256 hours, and the maximum time was 528. Within this framework the variation in the allowance for reading and writing ran from 129 to 312 hours, and for arithmetic from 63 to 118 hours.

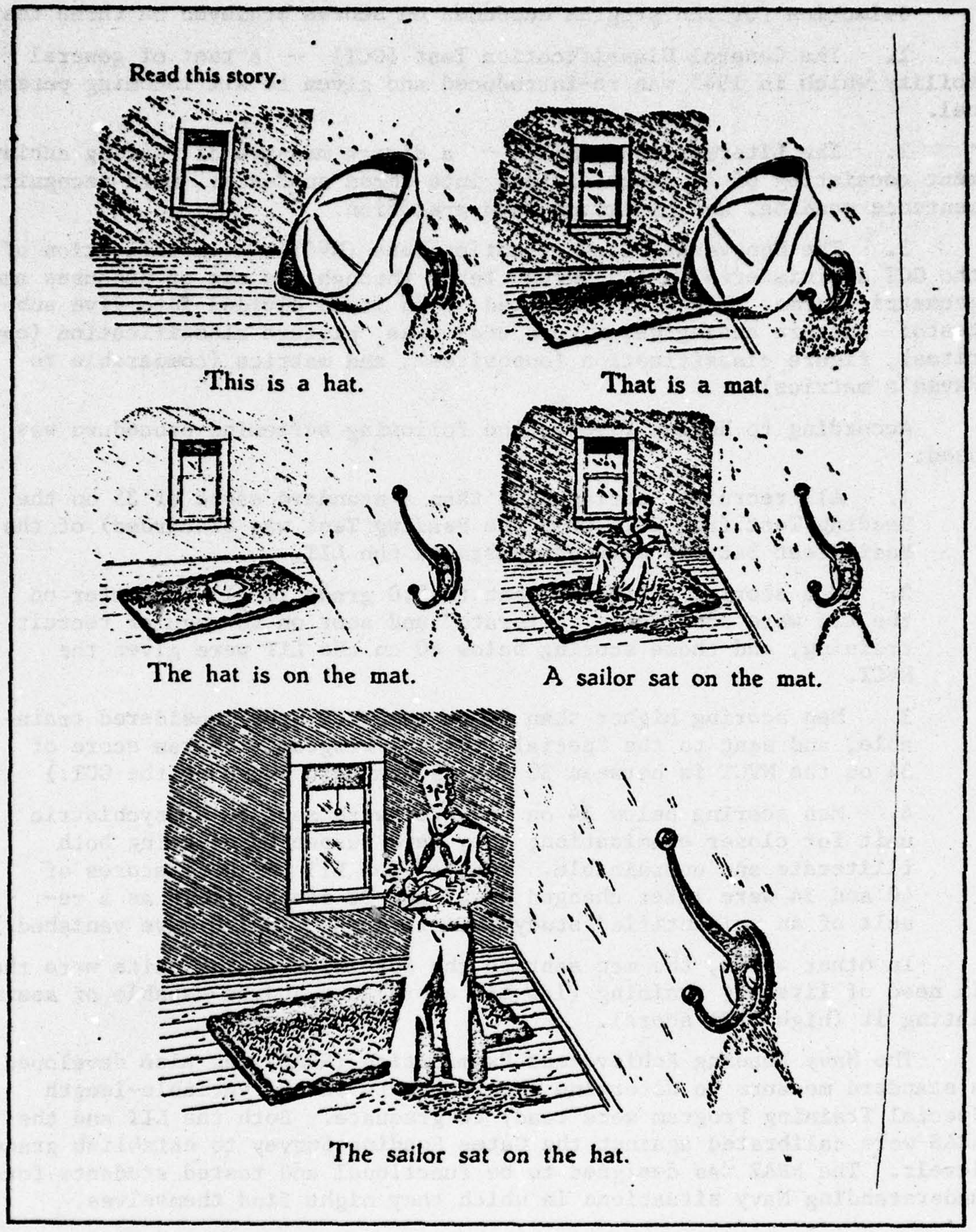


Figure 1. The first story in the *Navy Life* workbook series.

Selection for the program depended on scores achieved on three tests.

1. The General Classification Test (GCT) — a test of general ability which in 1945 was re-introduced and given to all incoming personnel.

2. The Literacy Test (LIT) — a direct measure of reading achievement consisting of 44 items divided into three subtests: word recognition, sentence reading, and paragraph interpretation.

3. The Non-Verbal Classification Test (NVCT) — a reflection of the GCT administered in non-verbal terms through the use of pictures and geometric forms. The NVCT consisted of 75 items divided into five subtests: picture analogies, figure analogies, picture classification (opposites), figure classification (opposites), and matrics (comparable to Raven's matrics).

According to Hodges (1964), the following screening procedure was used:

1. All recruits scoring less than a standard score of 35 on the Reading Test (the GCT, when the Reading Test was discarded) of the Basic Test Battery were administered the LIT.

2. Men scoring 40 (equivalent to 5.0 grade level) or better on the LIT were considered "literate" and sent on to regular recruit training, and those scoring below 40 on the LIT were given the NVCT.

3. Men scoring higher than 34 on the NVCT were considered trainable, and sent to the Special Training Program. (A raw score of 34 on the NVCT is between 30 and 35 Standard Score on the GCT.)

4. Men scoring below 34 on the NVCT were sent to a psychiatric unit for closer examination; they were suspected of being both illiterate and untrainable. (NB — The LIT and NVCT scores of 40 and 34 were later changed to 37 and 38 respectively as a result of an unidentified study, all records of which have vanished.)

In other words, the men sent to the special training units were those in need of literacy training (low LIT score) and judged capable of assimilating it (high NVCT score).

The Navy Reading Achievement Examination (NRAE) was also developed as a standard measure to determine when students in the variable-length Special Training Program were ready to graduate. Both the LIT and the NRAE were calibrated against the Gates Reading Survey to establish grade levels. The NRAE was designed to be functional and tested students for understanding Navy situations in which they might find themselves.

There appears to be a single evaluation study of the Navy's World War II Special Training Program. This study was administered and documented by Hagen and Thorndike (1953) on the basis of personnel records salvaged from the Naval Records Management Center, Garden City, New York. Two types of data were abstracted: (1) background facts that might predict success in the Navy, and (2) facts about the man's career that might indicate his success in Navy duty assignments. Three groups of men were identified:

1. The "Illiterate" Group. This group consisted of 1026 men inducted into the Navy during August and September 1944 and initially assigned to the Special Training Program at Camp Peary.

2. The Control Group. This group was selected by taking a serial number that was five digits above the serial number of a member of the illiterate group. The man thus identified was included in the control group if he was not black, not illiterate, and came from the same geographical area as his illiterate counterpart. If a serial number five digits higher did not identify a qualified counterpart, one five digits below was selected, and so on. In this way a control group of 1021 men was chosen. This group resembled the illiterate group on most geographical and demographic measures.

3. The Marginal Group. This group was composed of 999 men who had scored below 36 on the GCT and who entered the Navy at about the same time as the illiterates. This group was geographically and demographically dissimilar from both the illiterate and control groups.

Hagen and Thorndike described the average member of their illiterate group as follows:

> He was inducted in the Navy when he was about 19 years old. At the time of induction, he was single but had one or more people who were partially or wholly dependent on him for support. Before induction, he had lived in a rural area of the South where the standard of living was below the average for the nation as a whole. He had completed the fourth grade in school and left school at the age of fourteen after having repeated at least three grades. Since leaving school, he had worked for his parents or a relative on a non-mechanized farm. In his spare time he hunted or fished. He made a score of 4 out of a possible 17 on the Qualifications Test and a score of 31 on the Navy General Classification Test.
>
> (Hagen & Thorndike, 1953, p. 18.)

Hagen and Thorndike summarized the differences they found between the illiterate and control groups as follows:

1. The illiterates were much more likely than were the control cases to be assigned to construction battalions, and were less likely to be assigned to U.S. permanent party or to auxiliary ships.

2. The illiterates tended to receive a lower average proficiency in rate. Only 50% received an ave ge of 3.5 or over, as compared with 73% in the control group.

3. The illiterates received fewer promotions. Only 15% made petty officer, as compared with 37% in the control group.

4. The illiterates received more disciplinary actions. In the illiterate group, 23% had records of some type of disciplinary action, as compared with 11% of the control group. General courts martial were ten times as frequent in the illiterate group.

5. The illiterates more frequently lost time due to misconduct — 20% versus 7%.

6. The illiterates less frequently received an honorable discharge — 83% versus 88%.

7. The illiterates were somewhat more likely to receive a medication survey, and were the only group to be surveyed for inability to learn — 19% versus 15%.

8. The illiterates were somewhat more likely to incur a venereal infection — 5% versus 3%.

9. The illiterates were slightly more likely to generate a Veterans Administration disability claim — 11% versus 9%.

In most of the factors that distinguished the illiterate from the control group, the marginal group occupies an intermediate position, usually nearer the illiterate group than the control group. The only exception to this is the case of disciplinary actions; the marginal group were more often in trouble and their offenses were more serious.

The Hagen and Thorndike study is intersting, but it is essentially a study of personnel at different mental levels. It does not have much to say about the effectiveness of literacy training. It does, however, tell something about the success of illiterates in the Navy. In general, it seems reasonable to conclude with the authors that many or most of the illiterates appeared to make an acceptable adjustment to the Navy.

After the war, screening of illiterates with the GCT, LIT, and NVCT continued, but the Special Training Program was discontinued in the course of general demobilization. Ginzberg and Bray (1953) estimated that 35,000 men were assigned to the Navy's Special Training Program in the course of its history.

2. Recruit Preparatory Training

In 1950, with the influx of volunteers during the Korean War, commands were once more authorized and encouraged to identify and train any illiterates who might be aboard. By the spring of 1951, literacy training was being conducted informally in after-hours programs at the three recruit training centers (Bainbridge, Maryland; Great Lakes, Illinois; and San Diego, California). By September 1951, literacy training was planned on a full-time basis under the new title Recruit Preparatory Training (RPT) (Standlee, 1954). Officially, the main objective of the program was "to teach selected recruits to read and understand instructions and to prepare them to absorb military training. Recruits should attain a level of proficiency in reading which is comparable to completion of the fourth grade" (Curriculum for Recruit Preparatory Training, 1953, p. 5). Further, an accepting classroom atmosphere was to be created as indicated by the following guidance provided by the Bureau of Personnel.

1. Make the first reading tasks simple enough to insure that the recruit experiences an early feeling of success in learning to read.

2. Show the recruit that his low reading achievement is a handicap which must be overcome in reaching his goal — becoming a useful sailor.

3. Create a permissive classroom atmosphere and start where the learner is. The individual recruit is the one who has to learn. The most the instructor can do is guide and help him. The instructor should guard against a highly directive, subject-matter-centered approach which is on a level that never reaches some recruits.

4. Integration of military subjects with reading, writing, and arithmetic. Some examples of how this can be done are listed below:

   a. The instructor includes some topics related to service in the Navy at appropriate times in his reading, writing, and arithmetic classes. This procedure will probably be more effective during the latter part of RPT.

   b. Use of visits to ships or parts of the training center. Trainees can write about what they see, read accounts to each other, learn naval terms, etc.

   c. Learn to read safety signs found on the center or aboard ship.

   d. Preparation of RPT news sheet and use of local newspapers and announcements of recreation.

e. Provide guidance in writing letters home about barracks, field-day, bunks, meals, training activities, et cetera.

f. Familiarization with types of material found in *The Bluejackets' Manual*, libraries on the center and aboard ship, and courses available through USAFI.

(Curriculum for Recruit Preparatory Training, 1953, p. 11.)

The curriculum materials were, with minor updating, composed of the *Navy Life* series prepared for the World War II Special Training Program, although use of other supplementary materials was encouraged (Standlee, Fattu, & Auble, 1954a). The selection procedure for RPT was the following, as described by Hodges (1964):

1. All men scoring below 36 on the GCT were administered both the NVCT and LIT.

2. Men scoring 38 and above on both tests were considered literate and trainable and sent to regular recruit training.

3. Men scoring 37 and below on the LIT (38 on the LIT is about a 4.7 grade level) and scoring 38 or higher on the NVCT, were considered illiterate but trainable and sent to special training.

4. Men scoring less than 38 on the NVCT and on the LIT were suspected to be untrainable and were sent to a neuropsychiatric unit for further examination. From the neuropsychiatric unit, men were sent to special training, regular recruit training (on a trial basis), or discharged.

Referrals could also be made to the Special Training Program by Company Commanders in regular recruit training. The NRAE, with a new form added, was restored as an end-of-course criterion. The 5.0 grade level was still the course objective. Students were given 7-9 weeks to complete the course. An observer who was sent to the three RPT centers in 1953 reported that, in general, the selection and curriculum procedures in use followed the Bureau's directives, that a total of 239 trainees were enrolled in all three centers, that 85 to 94 percent of the trainees successfully completed the RPT program, and that they took 6-7 to 9 weeks to complete it (Fattu, Fay, D'Amico, & Standlee, 1954).

Two evaluative studies of RPT were completed. A study by Cofer (1954) investigated the effectiveness of RPT in achieving its most immediate objective, preparing the trainee for regular recruit training. Standlee (1954) dealt with the more ultimate criterion, performance of RPT graduates in fleet assignments.

A principal concept in Cofer's study is that of adjustment; specifically, adjustment of RPT graduates to regular recruit training. For the purposes of his study, mental health considerations were set aside in favor of adjustment that emphasized, first, adequacy of performance of recruit duties, and, second, attitudinal and motivational factors in relationship to the Navy, recruit training, and civilian plans. Fourteen rating scales were developed to assess adjustment in the first sense, and an 80-item questionnaire was developed to assess adjustment in the second sense. Two groups of subjects were identified; one at an early stage of recruit training (Group I), and one at a later stage of recruit training (Group III). Within each of these groups, four sub-groups were identified.

1.	RPT Graduates:	Group I	- N = 30
		Group II	- N = 21
2.	Recruits with GCT scores below 35, but not RPT Graduates:	Group I	- N = 13
		Group II	- N = 18
3.	Recruits with GCT scores between 35 and 40 inclusive:	Group I	- N = 29
		Group II	- N = 46
4.	Recruits with GCT scores above 40:	Group I	- N = 30
		Group II	- N = 49

Each subject was evaluated on his adjustment to regular recruit training by peers and Company Commanders separately on the 14 scales reflecting adequecy of performance in recruit training.

In general, members of the RPT sub-group resembled the members of the high GCT sub-group on the criterion scales more than they did members of the intermediate sub-groups. The results of the peer ratings for Group I subjects were the following:

1. The high GCT sub-group and the RPT sub-group were rated significantly superior to the intermediate sub-groups on personal cleanliness, military bearing, bunk and living quarters, general Navy performance, and leadership.
2. The RPT sub-group was rated as significantly superior to the other three sub-groups on marching and maneuvers and physical drill with rifle.
3. The high GCT sub-group was rated as significantly superior to the other three sub-groups on quickness to learn and response to orders.
4. The RPT sub-group was superior to the intermediate sub-groups on care of clothing.

The results of Company Commanders ratings of the Group I subjects were the following:

1. The high GCT sub-group was significantly superior to the other three sub-groups on quickness to learn, response to orders, willingness to work, and general Navy performance.

2. The high GCT sub-group and the RPT sub-group showed superior ratings on care of clothing and military bearing.

The pattern of ratings for Group II was essentially the same as that obtained for Group I.

In general, it seems reasonable to conclude with Cofer that RPT aided recruits in adjusting to regular recruit training, but it is still difficult to determine what contributions the cognitive aspects of literacy training made to this adjustment.

Standlee (1954) studied three groups of recruits in his investigation:

1.	RPT Graduates:	N = 611
2.	Marginal recruits with a GCT of 35 or lower, an NVCT of 38 or higher, and a LIT of 38 or higher:	N = 2,414
3.	Typical recruits with a GCT of 36 or higher:	N = 1,998

The 5,312 total subjects were originally selected by identifying all RPT graduates and marginal recruits processed during the period 10 May 1952 to 1 September 1952 and a random 10 percent of the typical recruits processed during the same period. Data on the fleet performance of these recruits were obtained by mailing out questionnaires to their Commanding Officers during July-August 1953. Most of the questionnaires were returned, of which, 5,023 were suitable for analysis — an effective response rate of almost 95 percent.

The RPT and marginal groups in comparison with the typical group were much more often assigned to general duty (unskilled work) and to duties that did not require as much reading skill. The two groups were less favored with promotions, had lower average performance evaluations by supervisors and Commanding Officers, and had a higher proportion of disciplinary actions and days lost due to misconduct or sickness. In the opinion of division-leading Petty Officers, the two lower groups generally showed less promise for future advancement and profitable service to the Navy. The two lower groups, however, indicated more intention to reenlist than the "typical" group.

Overall, the two lower groups tended to resemble each other, and to be somewhat less effective in performance than the "typicals". Where differences between the two lower groups were found, they tended to show the marginal group to be superior to the RPT group. However, most of the differences between the three groups were small and none of the groups had characteristically unacceptable performance; the majority of members in all the groups were judged to be serving the Navy adequately.

Two other relevant investigations came out of the general contractual effort that sponsored the Cofer and Standlee studies. Standlee, Fattu, and Auble (1954b) investigated the quality of Navy technical writing by tabulating the frequency of words appearing in the *Bluejackets' Manual* (14th edition), *This is Your Navy*, *All Hands*, and the *Naval Training Bulletin*; and Fattu and Standlee (1954) applied two Flesch readability formulas to the following publications: *Bluejackets' Manual*, *Stewardsman* (Manual), *Fireman* (Manual), *Steward and Cook* (Manual), *Commissaryman* (Manual), *This Is Your Navy*, *All Hands*, and *Naval Training Bulletin*. It seems unlikely that these investigations represent the first attempts to systematically judge the quality of technical writing in the Navy, but these are the first that appear to be available in the official literature. The results of these investigations are predictable: Standlee et al. showed that even experienced writers and teachers use too many rare words and exclude too many common words when relying on their experience and common sense; Fattu and Standlee found that standard, essential Navy publications were, most probably, too difficult for Navy enlisted men — the average Flesch readability score of their sample was 61.7, which converts roughly to a ninth-grade reading level.

With the end of the Korean War, sufficient manpower again became available to the Navy, and in 1957 the RPT program was discontinued in favor of a shift in research and administration emphasis toward problems connected with higher-level personnel. Subjectively, many persons in the Navy (including most of the RPT graduates) felt that RPT was well worth the time and effort; objectively, the evaluation findings were inconclusive.

3. Academic Remedial Training in the Present

The Navy's current excursion into literacy training appears to have been brought about by the Army's manpower supply problems during the Vietnam conflict. During the Vietnam buildup, the Army was forced to accept personnel who were classified as marginal by their scores on the Armed Forces Qualification Test (AFQT). Generally, these volunteers and draftees ranked in the 10th-30th percentile range on the AFQT (Project One Hundred Thousand, 1969). As a consequence of its manpower needs, the Army indicated to the Department of Defense that a coherent program for processing marginal personnel was needed if manpower requirements were to continue being met. Defense responded in October 1966, by establishing "Project 100,000" which was to help meet manpower supply problems by spreading marginal personnel throughout all three services. The project was also intended to provide training for these men so that they would be better prepared to return to civilian life if they chose to do so. Under Project 100,000, the Navy agreed to accept up to 15% of its enlistees from the marginal category and allow them to volunteer for the normal draft tour of two years. Accordingly, the Navy established RPT units at Great Lakes, Illinois, and San Diego, California, in February 1967 (Weeden, 1975). Almost immediately, the name of the program was changed to Academic Remedial Training (ART), and this is the name currently in use.

Despite the expiration of Project 100,000 in 1972, the ART program has continued in operation up to the present. Although projections have indicated that induction of marginal personnel will be unnecessary even under the current no-draft policy (e.g., Battelle et al., 1973), ART may be continued because of the rising national consciousness concerning equal employment opportunity and/or because of (anticipated) improvements in the national economy.

A student may currently be selected for ART if he averages between 3.0 and 5.5 grade levels on the Vocabulary and Comprehension Sections of the Gates-MacGinitie Reading Test. Originally, the *Navy Life* series was once more dusted off and re-edited for the program; however, use of these materials was largely discontinued in 1969. Weeden (1975) describes the new program that is projected for implementation in ART. The course of instruction may last a maximum of 8 weeks and is broken into two phases: a diagnostic phase and a therapeutic phase. The diagnostic phase provides for initial assignment of the student to one of three courses:

1. Course Mike — A phonics course emphasizing word attach skills.
2. Course Oscar — A reading course beginning at the 3.0 grade level.
3. Course Victor — A reading course beginning at the 4.0 grade level.

The therapeutic phase consists of two mutually reinforcing areas: word attack and reading abilities. The word attack area concerns phonics, vocabulary development, and word knowledge (root words, inflections, prefixes, synonyms, etc.). The reading abilities area concerns reading practice, comprehension skills, and study skills.

Notably, ART is integrated only into recruit training, as were RPT and the Special Training Program before it. The proportion of illiterates in the United States in 1960 was estimated to be only 2.4 percent (Harman, 1970), and, with current screening practices, it seems unlikely that there is an appreciable number of illiterates assigned to Navy billets. However, a new concern that keys on the concept of functional literacy is beginning to appear. It may well be that, despite the universal attainment of literacy by Navy personnel, some men (and women) may fail to perform because they do not read well enough to meet the requirements of their jobs; there may exist, in effect, a literacy gap. This possibility was first raised by the Fattu and Standlee report of 1954; however, two reports by Carver (1973a, 1973b) and current work by Duffy at the Navy Personnel Research and Development Center seem to support the current concern over a literacy gap. If such a gap exists, it may become necessary to extend literacy training from the recruit commands to the fleet. The nature and location of programs devised to meet the literacy gap would represent very new directions in the history of literacy training in the Navy.

C. Final Comment

From the preceding survey of literary training, it is apparent that problems of literacy are not a recent discovery; they have been a concern of the Navy throughout its history. This concern has been motivated both by a need for proficient job performance and by an interest in the general welfare of Navy personnel. In the development of literacy programs, the Navy has, at one time or another, been sensitive to at least the following six issues:

1. Level of Literacy. Efforts have been made to identify levels of literacy and to tailor instruction to the level required of a trainee. The idea of functional literacy as a consideration in determining what level of literacy a trainee ought to achieve in order to perform a specific job has also appeared.

2. Cause of Literacy Problems. Adjustments in Navy programs have been made to identify speakers from non-standard English linguistic communities, and to adjust instruction based on the language habits likely to be fostered by those communities.

3. Evaluation. Both narrow (Are program objectives being met?) and broad (Is fleet performance being enhanced?) issues have been considered in evaluating Navy literacy training.

4. Decoding. Issues of relating orthography to acoustic representation so that learners may identify information within their linguistic experience have been addressed.

5. Affect. Learners' attitudes toward themselves and toward literacy have been taken into account.

6. Active Participation. Most military training attempts to involve as much "hands-on" experience as possible and this approach appears in Navy literacy training particularly with the emphasis on the *Navy Life* workbooks.

On the other hand, the Navy's interest in literacy has been largely intermittent. From the evangelical concern of the Chaplains to the current interest in ART, lack of literacy skills on the part of Navy personnel has always been given full opportunity to handicap the Navy's operational effectiveness before anything is done about it. Programs that are devised in response to literacy problems usually prove to be relatively sophisticated, as the Special Training Program and RPT curriculums demonstrate, but these programs seem to appear only in response to wars or other national emergencies. The outlook for the current ART program cannot be very bright. ART is remarkable for continuing as long as it has after the demise of Project 100,000, but it seems reasonable to anticipate waning interest in ART on the part of the Navy unless problems arising from the all-volunteer military suddenly become more severe, manpower supply, for whatever reason, becomes more constrained, or the national priorities for equal employment opportunities are sustained.

The primary hope for continued interest in literacy training seems to rest on the concept of functional literacy in which the concern is not so much with illiteracy as with what literacy is necessary for specific Navy jobs (e.g., Duffy, Carter, Fletcher, & Aiken, 1975). In this respect it seems likely that concerns of literacy will be modified in two ways. First, they will become more narrow in that finding and applying technical information will be emphasized rather than more general skills of reading and writing. Secondly, these concerns will become more broad in that considerations of media, such as the organization and quality of technical manuals, the availability and capacity of systems for computer-based information retrieval, and the comprehensibility and useability of graphic and pictorial information will be integrated with more conventional literacy considerations. These are fairly novel directives for literacy training in the Navy, but they are also timely. It is difficult to come away from an historical survey of literacy training without an appreciation for the continuing increases in the amount of information that must be available for successful performance of Navy jobs. Classically exponential, these increases were at first gradual, but are now accelerating at an impressive rate. The importance, and even urgency, for systematic concern with literacy is hard to escape, and it is to be hoped that literacy training will receive continuing support from all the military services.

COMMENTS ON THE PAPER BY FLETCHER

Robert Glaser
University of Pittsburgh

Glaser based his remarks on Fletcher's paper and on discussions held at the Conference up to that time. His major point was that in both civilian and military settings, there seems to be a shift from speaking about general literacy and general intelligence toward the consideration of specific kinds of literacy and specific intelligences. He identified *functional job literacy* as one type of alternative to the concept of *general literacy*. With *functional job* literacy, the concern is to turn someone into an expert on his job — to give someone the wherewithal to develop expertise in his job and to develop pride in this job. He suggested that we frequently confuse functional literacy with general literacy, rather than recognizing it as the knowledge structure of job tasks and the interrelationships of fundamental concepts in jobs. Functional job literacy training should provide a knowledge network so people can retrieve and use job information readily, easily, and with a high degree of connectivity to other concepts so that this information can be used most adaptably. He mentioned the primitive theories of knowledge and conceptual networks now being thought about in cognitive phychology, and suggested they might be valuable areas for research related to functional literacy. In regard to *general literacy*, Glaser stated that rather than regarding general literacy in terms of 4th, or 6th, etc. grade levels, we should consider general literacy in terms of the number of opportunities open to a person with certain literacy skills. He related the concept of aptitude to that of general literacy by pointing out that we have not really investigated the trainability of what we call aptitudes, and similarly, the grade level concept of general literacy does not focus on the trainability of literacy. There are a finite set of skills needed to get through school in the way we currently train people, and literacy skills make up a heavy component of this training. The more we use the printed page, the more certain skills are required to get through the training program; this certain set of skills then becomes rarefied in what we call "general literacy". If we want to maximize the number of people getting through training, we have to analyze the basic components and processess making up general literacy and general intelligence and train these as specific skills. As one way to train general literacy, Glaser suggested first training people in functional literacy; this might then provide the motivation to improve their general literacy. He also recalled the need to think about other ways of learning by non-literacy skills, and cautioned about "over-literacizing" a job; that is, making the job more "prestigefull" by an unnecessary proliferation of complicated printed materials.

A HISTORICAL PERSPECTIVE OF ARMY R&D IN READING AND TECHNICAL WRITING

Jack H. Hiller

U.S. Army Research Institute

Army use of technical writing for training and operations has been and continues to be extensive. The Army has trained at least a half million illiterates since 1942 to help them to read this literature and to enhance their adjustment to Army life. Yet, research on reading and technical writing was sparse. I will attempt to characterize this R&D in reading training and technical writing from the Second World War to the early period of U.S. combat in Vietnam. My purpose is to provide a backdrop against which contemporary R&D may be compared to reveal both progress and research gaps.

Reading Training

There were two major literacy (reading, writing, and arithmetic) programs conducted during the time frame for this review: (1) the Special Training Unit (STU) program of WW II (Goldberg, 1951), and (2) the Basic Education Project at Fort Leonard Wood during 1953-4 (Goffard, 1956). These two programs provide the focus for this review, and they will be discussed together as we proceed. It should be noted that the STU program met an urgent need and was not experimental, while the Fort Wood project was.

This review is organized according to a systems approach toward training so that the following topics will be covered: job requirements, training goals and objectives, instructional methods and media, selection and training of instructors, selection of students and achievement testing, and program evaluation.

1. Job Requirements. Formal analyses and measurement of general Army reading needs were not conducted. Likewise, measurement of reading needs posed by specific jobs and job training was not performed. Informal, intuitive estimation of vocabulary and reading tasks provided the basis for training.

2. Training Goals and Objectives. In both programs, enhanced trainability and military performance were primary goals. The STUs were furnished with relatively clear and complete learning objectives that appeared to be functional to Army needs. For example, the first objective required recognition and understanding of 46 specific Army words. Other objectives were to follow directions and to find and retrieve specific information. The STU reading materials were carefully selected to be directly relevant to Army life. In contrast, the reading goals and materials of the Fort Wood project were academically oriented.

3. <u>Instructional Methods and Media</u>. The STUs were provided with numerous training aids, such as film strips, flash cards, posters, and charts, as well as supplementary periodicals. The core of training was a specially prepared basic text and workbook that was graduated for difficulty. All materials were Army relevant. The Fort Wood project used commercially prepared materials.

4. <u>Selection and Training of Instructors</u>. Mainly experienced, well-educated teachers, civilian and military, were employed by the STU. Supervisors and teachers were given pre-service and in-service training. Carefully prepared instructor guides based on STU experience were provided. Only civilian instructors worked at Fort Wood.

5. <u>Selection of Students and Achievement Testing</u>. Students sent to the STU were drawn from three categories: (1) illiterate, but able to learn, (2) non-English speakers with learning potential, and (3) men from the lowest category of learning ability. Screening was accomplished by a series of tests. Individuals who passed a specially developed literacy test were given the Army General Classification Test. Of this group, the lowest performers were sent to the STU. Individuals who did not pass the literacy screen were given one or more non-language ability tests to estimate learning potential. Low scorers on the non-language ability tests were discharged.

Selection at Fort Wood was based on having a relatively low aptitude area score on the Army General Classification Test and a score indicating reading ability level below the fourth grade on the U.S. Armed Forces Intermediate Achievement Test. This achievement test is based on a commercial test for elementary school children and covers vocabulary, reading comprehension, and arithmetic. Both the vocabulary choices, and skills reflected in comprehension test items appear to have marginal relevance to Army needs.

The STU attempted to place individuals according to a test specifically designed for that purpose; however, placement was intuitively decided and definitive research was not conducted. A terminal achievement test, having the same form as the placement test, was used to help determine the decision to graduate students. These tests relied on content validity and internal consistency to establish validity. In addition, they were correlated with a commercial elementary school test to support validity and to estimate grade equivalence. Time allowed for graduation from the STU varied with Army manpower needs, but overall the maximum allowed was about 12 weeks. Graduation from the Fort Wood program required achievement of the fourth-grade level within 96 hours or 16 days of classroom instruction.

6. <u>Program Evaluation</u>. For a brief period during the early days of the STU, questionnaires were sent to training units who received STU graduates. The training units found these questionnaires burdensome, so the War Department banned any further inquiries on STU post-graduation performance. Evaluation of the STU therefore, must rely on criteria

internal to the program. Attributing effects to the STU would have been problematical anyway; since its students were heterogeneous, instructional procedures, while carefully supervised, allowed for local innovation, and graduation was based on a multitude of criteria that kept changing. In addition, control Ss were not employed.

Effectiveness of the various instructional programs conducted at Fort Wood was evaluated against external performance criteria, with the aid of specifically selected control subjects. Higher scores were reliably achieved by the experimental subjects on both written and performance tests concerning basic military proficiency. These military tests were given to all trainees, including graduates of the Fort Wood literacy classes, at the eighth week of basic military training. However, the gain associated with the literacy training was only 2-3% and was regarded as practically unimportant. Supervisor ratings based on six months at the soldier's first duty position showed no practical effects. However, these negative results must be regarded as equivocal, since the graduation test was based on elementary school-related rather than Army-related achievement, and the criterion level of fourth grade may well have been too low to have impact.

Summary

Army R&D on reading training up to 1954 did not establish clear evidence of what reading needs exist and what effects reading training may have on military training and performance.

Conclusion

From the perspective of a systems framework, what little research was conducted cannot be regarded as in any way conclusive. Specific reading needs were not researched. Training and job requirements were not specifically modified to avoid or minimize literacy demands (although there was debate on this topic in the Departments of War and Army). Alternative methods of instruction were not extensively evaluated. The graduation criterion was set at fourth grade, which might be too low. In the largest training program, that of the STU, formal data concerning the effects of literacy training on Army performance were not collected. Research form the Second World War through the Korean War may be characterized as aiming at relatively low skill levels using brief training programs. Such brief training reflects a determined effort to provide training that is cost-effective, although research to estimate costs and effects has been too limited to guide specification of training parameters.

Research on Technical Writing and Useability

This review is organized according to a system engineering approach, with attention given to the following selected topics:

- Determination or specification of job functions and tasks performed by technical writers.
- Determination of skills and knowledge required for training and job performance by technical writers.
- Development of guidelines, specifications, and acceptance criteria for training materials, technical documentation, and job aids to be produced by technical writers.
- Evaluation of training materials and job aids (e.g., writer's handbook) for technical writers.
- Comparative evaluation of the effectiveness of materials produced by writers having different training or using different job aids.

1. Job Analysis. I have not found any evidence of systematic research to determine the job functions and tasks performed by individuals who were assigned technical writing responsibilities. Apparently, as Richard Kern found by interviews conducted in the early seventies (reported in this volume), technical writing assignments were made on the basis of job proximity, with subject matter familiarity presumed to provide the necessary and sufficient expertise. Given this presumption, it follows that research into job requirements would not have been pursued.

2. Skill and Knowledge Requirements. Although empirical research aimed at determining tasks performed by technical writers seems not to have been conducted, it is clear that the communicability of Army writing was widely recognized to be a problem. Nevertheless, specific skill and knowledge requirements were not researched.

3. Writing Guidelines and Acceptability Criteria. To alleviate problems caused by needlessly difficult writing styles, various writing guides were compiled; for example, Klein (1946), Stephenson (1950), Milton (1953), Hoehn (1960), and Rogers and Thorne (1965). Klein presented a guide for writing military correspondence, orders, reports, manuals, and news copy. He stressed that clarity and conciseness were primary virtues. To help writers achieve clarity, he showed examples of writing with specific faults and corrected versions. He also identified confusing military jargon, and provided brief rules of grammar and some suggestions for writing manuals.

Milton prescribed the use of topic sentences, transitional techniques, "concrete" words, active verbs, and familiar vocabulary; and the avoidance of passive verbs, parenthetical expressions, and other complex sentence forms. He also advocated the Flesch notions on using short sentences with

personal pronouns. Stephenson reported that the Infantry School at Fort Benning had adopted a Flesch (1948) readability formula to rewrite its infantry manuals.

Hoehn described certain logical criteria for deciding when information should be conveyed by a handbook instead of by training:

a. Ease of Communication. Consider using a handbook if it is easy to show how to perform a given task.

b. Difficulty to Learn. If a task is difficult to learn (memorize), but can be effectively supported by a handbook, then skip the training in favor of the handbook, or support performance with a handbook.

c. Frequency of Use. If a task is seldom performed, include it in the handbook rather than training.

d. Accuracy or Safety Requirements. "In situations in which the safety of the repairman or the effectiveness and safety of the equipment allows for no errors in performance, responses should be supported by handbook content as well as training." (page 6).

Hoehn also described certain logical criteria for producing or selecting handbooks: (a) All of the steps required to perform a given task should be covered, (b) the handbook should be easy to use, (c) the handbook should be cost-effective by either improving job performance or reducing training costs.

Included in Hoehn's report was an exemplar for handbook design in the area of electronics troubleshooting. However, one is inclined to wonder if an "average" Army writer could readily understand and apply this exemplar. Likewise, could even psychologically sophisticated writers use Hoehn's logical criteria; for example, could they accurately predict the communicability and learning difficulty of proposed handbook content, as compared to active training? Thus, it seems to be that the useability of Hoehn's logical criteria is a question.

Rogers and Thorne actively considered factors influencing the decision on what to include in a troubleshooting manual, and on how to organize and present its contents. Based on their theoretical analysis, an experimental manual was prepared. An experimental test was conducted by having one group of technicians use the experimental manual to locate malfunctions in comparison to a group of technicians who used standard materials, including their personal notes. The technicians using the experimental manual performed "substantially" faster and more effectively. Based on this research, a guide for publication agencies to use when preparing similar manuals was written (Rogers & Harris, 1964), but I have not found evidence that it was field tested.

Across the years, it is possible to see a trend from untested, logical prescriptions on how to write effectively (Klein, 1946) to empirical testing of hypotheses on how to present information (e.g., Rogers & Thorne, 1965).

4. Evaluation of Training Materials and Job Aids for Technical Writers. I have not uncovered any instances where the utility of writing instruction or writer's manuals was empirically evaluated. At the very least, the actual or attempted useage of skills developed in training should be measured, as well as the useage of manuals on the job. If a writer's manual is never taken off the shelf, then research to design a better one might be seen as futile. When a manual is *looked* at, it may offer deceptively simple advice, such as "write short sentences using familiar vocabulary, action verbs, personal pronouns, and few prepositions", which is difficult for a writer to effect without further guidance or training. Furthermore, there is no solid evidence that writers who conform to prescribed rules actually produce more effective writing.

5. Comparative Evaluation of Alternative Training or Job Aids. It would seem to be a point of dubious finesse to mention that comparative evaluations do not appear to have been conducted, after having cited a lack of evaluation for singular products. However, it seems appropriate to emphasize that the ultimate objective for research on technical writing is to improve the readability and useability of technical documentation. Ideally, the products of writers who use a given experimental writer's manual would be compared for useability with the products of writers who are left to their own devices, or who use an alternative job aid. Unfortunately, such useability research appears to be difficult and expensive (see Richard Kern's paper in this volume).

OVERALL CONCLUSION

The major impression of research on reading and technical writing gotten by this reviewer is that, as of the mid-sixties, there was a paucity of research. One might speculate on the reasons. On the one hand, research to improve communication has faced the spectre of theoretical and empirical ignorance about how communication skills are learned and can be taught. On the other hand, motivation to embark on potentially costly research programs aimed at improving training and job performance by increasing communication skills appears to have been relatively low, as compared to the motivation to acquire new hardware.

COMMENTS ON THE PAPER BY HILLER

Jeanne Chall
Harvard University

Chall made five points relating not only to Hiller's paper, but to other presentations and discussions. The first point was that we need to keep in mind the fact that concerns for readability measurement and the writing of readable materials are not restricted only to the production of low level materials for beginning or low level readers. These are also concerns for the production of higher level materials, such as newspapers, textbooks, corporation reports, and so forth. Yet, the problem of readability measurement and writing readable materials may be quite different for providing materials for developing readers than those for highly skilled readers. She related this to the expressed concern for functional literacy (see Guthrie's comments on Burkett's paper) in which the primers for the illiterates incorporated meaningful content, and contrasted this with the type of primers constructed to teach basic phonics skills. The questions of whether one or the other approach is best is open, Chall suggested.

The second and third points made by Chall were that more needs to be known about the unskilled readers who are entering the Armed Services. For instance, how many are of both low IQ *and* unable to read, and how many are of normal IQ but are dyslexic — unable to read, but otherwise intelligent. For a completely adequate literacy training program, one needs information about the probable reasons for a person's difficulties. Chall's fourth point concerned the objectives of literacy training programs within the Army and their practicality. She expressed her opinion that any program for adults which had only a fourth-grade or fifth-grade skill level as an objective was not preparing the person for any practical tasks. She wondered why it was considered possible that one could take somebody who goes to school for eight to twelve years and does not learn to read, and put that person through a 12- (or so) week program and teach him to do something (read) that a normal child would take four or five years of constant practice to achieve? Chall suggested that comparisons be made of adult illiterates in the military with civilian illiterates and with children who are learning to read at different ages. From such comparisons, military researchers might gain useful information about the aspects of learning to read that are universal to all and about those specific to adults, and further, to adults in the military. Where there is commonality, much can be learned from the extensive research that already exists, based on children.

The fifth and final point made by Chall was that there needs to be materials of a suitable reading level around for newly trained literates to practice reading. If Army men are taught to read at the fourth-grade level, and there are no Army materials written at that level, then there is very little opportunity for further practice and skills may dwindle to pretraining levels. Army and other military literacy training systems must be designed to promote the continued growth of newly acquired, fragile reading skills.

A REVIEW OF RESEARCH AND DEVELOPMENT ON LITERACY AND TECHNICAL WRITING IN THE AIR FORCE

James R. Burkett

Air Force Human Resources Laboratory

INTRODUCTION

Problems relating to reading, literacy, and technical writing have persisted in the Air Force for many years. Because of the impact of these problems on training, job performance, and operating costs, as well as the considerable complexity of the literacy area from both a men and a materials standpoint, numerous R&D efforts have been carried out addressing various aspects of literacy and technical writing problems in the Air Force. The purpose of this paper is to summarize the major research directions the Air Force has pursued in these areas over the years in order to provide some historical perspective on what has been accomplished and what has led up to the present R&D programs on literacy and technical materials.

Before reviewing the various Air Force research efforts concerned with the literacy area, some perspectives on the nature of the problems that have generated the research are in order. Simply stated, both literacy and technical writing problems arise from a situation where there is a mismatch (or discrepancy) between the reading demands of job or training materials and the reading abilities of personnel who must use those materials. This mismatch between the literacy skills of men and the readability or useability of their materials is often referred to as a "literacy gap". The existence of a literacy gap interferes with the successful transfer of information from written materials to the personnel who must use that information for the purpose of training or job performance. The Air Force has always had some numbers of lower or limited ability personnel and has also experienced an ever increasing technical complexity of job demands and materials, even in high ability career areas. Accordingly, a primary goal of Air Force researchers, personnel planners, and training managers over the years has been to identify and reduce mismatches between men and materials as they have occurred, and a variety of techniques and approaches toward that goal have been developed and employed with varying success. The important consideration here is that the problem has never been found to be the fault of the men or the materials alone. Both have contributed to the magnitude of the recurring literacy problems faced by the Air Force, and research has had to deal with both to some degree to develop the tools and techniques to address the problems that have persisted through the years.

Thus, the goals of improving both literacy skills and technical materials have proven to be complex and difficult problems for the Air Force, requiring recurring, if not altogether ongoing, research and development emphasis over the years, an emphasis which continues today in the current

program of the Air Force Human Resources Laboratory. The work accomplished throughout the history of Air Force research on the problems of literacy and technical writing spans a wide variety of topic areas and approaches. Major areas of emphasis have included efforts on readability, technical writing and formatting, procedures for dealing with lower ability personnel, job performance aids, useability of technical materials, simplification of materials, literacy skills and literacy training, determining reading requirements, and measuring textual comprehensibility. While the R&D efforts accomplished to date have provided a useful technology base and a variety of useful techniques and applications in these areas, the work continues. This is a reflection not of the irrelevance or lack of success of earlier efforts, but of the scope and complexity of the problem area and of the ever changing nature of Air Force jobs and manpower.

We shall now turn to a review of Air Force R&D on literacy and technical writing. First, early efforts will be described. Next, the work that emerged from programs of the Air Force Personnel and Training Research Center (AFPTRC) from the middle through the late fifties will be reviewed, along with the research relating to lower ability airmen that was continued by subsequent organizations through the middle sixties. Then, the history of the Air Force program on Job Performance Aids will be briefly described. Other research of interest that has occurred outside the main Air Force personnel, training, and human resources R&D channels will then be noted. Finally, recent research on literacy and readability will be mentioned to complete the historical perspective of R&D accomplished to the present time. Throughout the review, both Air Force in-house and Air Force sponsored contract research are included, if the work appeared relevant to the topic areas. While it is recognized that much of the Air Force R&D on literacy and technical writing has been accomplished by contractors, no attempt has been made in this review to separate the contract efforts from the in-house work.

EARLY EFFORTS ON READABILITY, TECHNICAL WRITING, AND LITERACY

In the early fifties, very soon after the establishment of the Air Force as a separate service department, interest in research on readability of job and training materials and research on utilization and training of low ability personnel emerged as important concerns. The Air Force was becoming aware of the fact that much of its job and training materials were very difficult to read, and that the increasing influx of lower ability airmen coming in as part of the Korean Conflict manpower buildup was compounding the problem. England (1950) used the Flesch (1948) readability formula to check Air Force publications and found them to be unnecessarily difficult. He advocated a campaign for use of "plain talk" and simplification of materials in accordance with Flesch's reading ease guidelines. A year later (England, 1951) he reported an analysis of the effects of a one-year application of his "plain talk" principles to USAF Air Materiel Command publications. Again, using the Flesch formula to

check the readability of his simplified publications, England found substantial gains in reading ease and comprehension among users of the simplified materials. He advocated their expanded use, but it would appear that his campaign for "plain talk" had only a limited and somewhat transient impact on the problem of too difficult job materials.

Training personnel were also becoming interested in the potential application of popular readability metrics for improving their instructional materials. Hegg and Weaver (1952) evaluated the reading difficulty of training materials at Vance AFB using the Flesch formula and found that the training literature sampled was written at a level far above the reading ability of the middle two-thirds of the students who had to use the materials. While they offered no suggestions for specific ways to simplify the training materials, they did indicate that instead of writing to a 12th- or higher grade level, writing to about a 9th-grade level would better match the reading capabilities of most of the trainees. Peerson (1953) reported on a readability improvement program for locally produced training materials at Keesler AFB. This time, the Dale-Chall (1948) readability approach was used, and again the idea was to reduce the difficulty of the training literature to a level more appropriate for the students. Milton (1953) suggested that technical writers should take the initiative in improving the readability of their materials by using a variety of available techniques (such as those advocated by Flesch) for making their writing less complex, more direct, and easier to read. The emerging interest in measuring and improving readability of Air Force writing was officially recognized in 1956 with the publication of a *Guide for Air Force Writing* (AFM 11-3). This manual gave instructions for calculating the Fog Count readability measure developed by McElroy around 1953 and recommended the application of a variety of effective writing principles, including those of Flesch and others, by Air Force personnel. While this guide was a useful step, the mere issuance of such a manual could not resolve the overall literacy and technical writing problem. For one thing, an increasing influx of lower ability personnel in the early fifties had resulted in training problems of sufficient magnitude to make literacy a major concern.

The Air Force had always had some low ability personnel (mental Category IV on the AFQT) and some effort had been directed toward reducing their training problems through special programs to delete difficult reading materials from their curriculum and to provide extended course lengths to give them more time in training (Horton, 1951). However, by 1952 the problem had become more acute, and the Air Force undertook a comprehensive series of studies of special training programs for limited aptitude airmen with marginal literacy skills. The first study in this series (Gragg, Kieselbach, Murphy, Peckham, & Heller, 1955) attempted to determine the effectiveness of covering the same material then in the standard 8-week basic training course (which included language arts) at a slower pace in an expanded 14-week course using matched groups of marginal airmen. The two training programs differed only in the length of training time provided. The performance of the two groups was compared at the end of their

programs on various measures, and no appreciable differences between them were found. There was no difference in language arts (as measured by the California Achievement Test) even though a major portion of the expanded time in the 14-week course had been devoted to that area. It was concluded that little gain could be expected for low ability airmen from just extending the training period and covering the standard material at a slower rate. A more extensive study was then undertaken which incorporated a larger sample of marginal airmen, two new experimental curriculums, and extensive field follow-up comparisons between the experimental low ability groups and control groups of higher ability men.

The new study was known as "The Air Force 'Project 1000' Study". The subjects were 1000 limited aptitude airmen. One-half of them received basic training in an experimental six-week curriculum and the other half in an experimental twelve-week curriculum. The two sample groups were matched on age, race, marital status, and prior education. The two training programs differed primarily in the inclusion of 45 hours of language arts, and 45 hours of mathematics training in the longer course, and in the expanding of instruction in military fundamentals from 18 hours in the six-week course to 35 hours for the twelve-week course. Mastropaolo, Carp, Erdmann, and Schmid (1954) reported the relative effectiveness of the two training programs on a variety of criteria obtained at the end of training and after six weeks of initial duty in the field. At the conclusion of training, the twelve-week group showed a small margin in arithmetic skills over the six-week group, but none in reading skills or language arts. When compared six weeks after assignment, no differences were found between the two groups on measures of job proficiency, attitude, or adjustment. A control group of "normal ability" airmen who had undergone the standard training course that were working alongside the low ability men were found to be superior on job performance ratings and achievement test scores, although somewhat inferior to the low ability groups on attitude and adjustment variables. It was concluded that for practical purposes, the longer course was not superior to the shorter one for the low aptitude airmen.

An eight-month-after-assignment field follow-up study was also conducted (Mastropaolo, Carp, & Erdmann, 1954) and the findings were generally consistent with those of the first study showing that the two training programs had produced equivalent results in terms of on-the-job performance, attitudes and adjustment. Again, a control group was found to be higher in job proficiency ratings and achievement test scores than the Project 1000 airmen. Thus, the results of the eight-month follow-up indicated that the low aptitude men had not benefited from expanding their training in terms of their job proficiency, retainability, or conformity to other Air Force standards of performance and discipline. It also appeared that those marginally literate men who had received additional training in language arts and mathematics during the longer course were not appreciably different in their skills, knowledges, and adjustment to the Air Force than

those marginal airmen who had not had such additional instruction. The value of including language arts instruction during basic training appeared questionable.

Shanley and Smith (1955) attempted an evaluation of the effectiveness of the language arts aspect of the basic training program for low ability airmen (Category IV on the AFQT) to determine if the procedures being used for remedial training in reading, writing, and spelling were having the desired impact. Three groups of 140 marginal airmen each underwent one of the following training programs: (1) 45 hours in language arts with military subject matter as the vehicle, (2) training in the same military subject matter but with no training in language arts, and (3) the regular non-remedial training. Scores of the groups on reading comprehension and military subject matter tests at the end of training indicated that the language arts program was effective in improving the comprehension of written material. However, it was also noted that there was great variability in the degree of improvement across trainees, and that the previous studies had shown basically negative results with an even more extensive literacy program. It was concluded that remedial literacy training could only be expected to show some benefit for marginal groups with a fairly wide range of ability, and that such programs might prove most useful as criteria for separating those men with a mental deficit from those with an educational deficit.

RESEARCH ON READABILITY, TECHNICAL MATERIALS, AND LITERACY SKILLS FROM THE MID-FIFTIES THROUGH THE MID-SIXTIES

By 1954, the Air Force Personnel and Training Research Center (AFPTRC) had been established, and this organization had a large-scale and diverse R&D program well underway. As will be noted later, the extensive Air Force R&D program on Job Performance Aids (JPAs) originated within AFPTRC during this period. A series of studies different from those specific to JPAs was carried out during this period by Klare and his associates as a result of the increasing interest in improving Air Force technical materials. This work involved an extensive series of experiments and studies published from 1954 to 1959 investigating the relationship of a wide variety of communication, textual, and readability variables to the comprehension, learning, and retention of technical materials.

In the initial report in this series, Klare, Mabry, and Gustafson (1954) present the results of a number of experiments on the effects of varying readability and level of abstraction, use of personal words, and underlining key words and phrases on the effectiveness of written technical training materials. Effectiveness was assessed by measuring immediate and delayed retention of the material, reading speed, interest value, and "acceptability" in terms of attitudes toward the material. Increasing readability (as measured by the Dale-Chall, 1948, and Flesch, 1948, formulas) was found to produce increased immediate retention and reading speed and the more readable materials were judged more interesting by the subjects. Use of personal rather than impersonal words did not appear to increase the interest or acceptability of the material.

AD-A034 730 HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA F/G 5/10
READING AND READABILITY RESEARCH IN THE ARMED SERVICES.(U)
SEP 76 T G STICHT, D W ZAPF N00014-76-C-0312
UNCLASSIFIED HUMRRO-FR-WD(CA)-76-4 NL

2 OF 4
AD-A
034 730

Klare, Mabry, and Gustafson (1955a) described their attempt to relate patterning (underlining) of technical materials to immediate retention, amount read in a given period of time, and acceptability of the material. Underlining important words resulted in more lines read and in higher retention than either underlining words which appeared on tests or no underlining, but the difference was not significant. It was further noted that patterning as used in this experiment appeared to have little impact on either reading speed or acceptability. Klare, Mabry, and Gustafson (1955b) further explored the question of holding content and style difficulty of technical material constant to determine if increasing human interest in such material would alter immediate retention, acceptability, and reading speed. Two levels of human interest were used, varying in both percentage of "personal words" and percentage of "personal sentences". The high level scored 28 ("interesting") on the Flesch formula, and the low level scored zero ("dull"). Results indicated that a high level of human interest of the technical writing produced no significant difference in retention and was consistently judged less acceptable. However, it did show a slight tendency to produce a greater amount read in a given time as compared to the low human interest level material.

Klare, Mabry, and Gustafson (1955c) further elaborated their findings on attempts to relate style difficulty to immediate retention, reading speed, and acceptability of technical training material. An easier style of writing (as measured by the standard formulas) was found to result in greater and more complete immediate retention, a greater amount read in a given time, and more acceptable reading material. This study also indicated a greater importance of content than style in determining how well material will be accepted by trainees. In addition, it showed a high relationship between judgements of material as easier to read and more pleasant to read. Another study in this series (Klare, Gustafson, Mabry, & Shuford, 1955) was conducted to explore the relationships of immediate retention test scores covering a technical training passage, selected "career preferences" in the form of expressed interests, and certain airmen aptitude indices. No significant relationship was found between career preferences and retention test scores, but a highly significant positive relationship was clearly indicated between aptitude scores and retention.

Building on the earlier studies in the series, Klare, Shuford, and Nichols (1957) examined the relationship between style difficulty of technical passages, practice in the form of additional readings, and aptitude to both reading efficiency and retention. Results indicated that an easy style produced significantly higher reading efficiency and retention, and that additional practice increased words read per second, recall, and word recognition. Higher ability airmen read more efficiently and with better retention than lower ability men. In another study in the series, Klare, Nichols, and Shuford (1957) assessed the effects of three types of typographic arrangements and long and short "thought units" on reading efficiency, acceptability, and immediate retention of technical material.

Results suggested that a square-span and a spaced-unit arrangement might possess some advantages over a standard sentence arrangement. The new arrangements were found to be less acceptable to the trainees than the standard sentence format, but this feeling was less when the "thought units" in the new arrangements were small rather than large.

Klare, Shuford, and Nichols (1958) reported a study of the relation of format organization to learning technical material. The use of heads and side heads in a text was tried as a learning aid, and the effect of changes in levels of organization of the material on its acceptability to trainees was also assessed. No significant increase in reading speed was found for higher levels of organization over lower levels, but trainees significantly favored the higher levels in terms of acceptability of material. Scores on an immediate retention test also appeared to favor higher levels of organization. Thus, this study did demonstrate an increase in reader acceptance for more highly organized as compared with less organized material. The question of whether the addition of headings increases immediate retention and comprehension was unresolved, but this study provided at least a qualified positive answer. In a study related to the previously described series, Stolurow and Newman (1959) attempted to isolate patterns of relationships among previously identified stylistic features of printed material in order to obtain a smaller number of more fundamental variables for research use. Gray and Leary's (1935) intercorrelation matrix was reduced from 44 to 23 variables and factor analyzed. After rotation, two factors tentatively described as easy versus difficult words and difficult versus easy sentences were found to account for approximately 50% of the variance, and the factor analysis provided useful new insight into the relative importance of these and other factors for readability measurement.

Finally, we shall consider an AFPTRC study not related to the previous series, in which Rubenstein and Aborn (1958) examined the interrelationship between learning, word-prediction, and readability. Correlations computed between amount learned, the number of correct predictions per word, and Flesch and Dale-Chall readability scores indicated the following relationships: Learning, prediction of words, and readability were closely interrelated; prediction and readability correlated about equally well with learning; the Dale-Chall formula correlated more closely with learning than the Flesch formula; and, despite the differences between the two readability formulas, they were found to correlate very highly with each other, which supported the findings of other researchers suggesting that the two formulas measured substantially the same things.

All of these studies of various aspects of readability and technical materials through the late fifties contributed to the beginnings of an Air Force technology base in the technical writing area which later efforts would build upon. Also during this period and into the mid-sixties, personnel researchers were actively engaged in studies and analyses of ability and educational requirements, personnel skills and characteristics, and other matters relating to low literacy and low aptitude airmen which the Air Force still had in limited numbers.

McReynolds (1958) noted that the quality of career enlisted personnel was up during the 1957-58 period, and that the rapidly increasing complexity of Air Force weapon systems would require additional high aptitude career personnel in coming years. Airmen in later enlistments were found to be typically higher in aptitude than first term airmen, and the need for continuing numbers of high ability personnel to man the highly technical career fields was noted. Flyer (1959) found that low educational level was the best single predictor of unsuitability discharge, and Judy (1960) examined educational requirements for various specialties and found that high school graduation was correspondingly the best single predictor of success in Air Force technical schools. Gordon and Bottenberg (1962) subsequently confirmed that of the variety of individual predictors of Air Force success, the amount of education attained was the most valid, further justifying the Air Force in limiting recruitment to high school graduates.

Lecznar (1962) traced the aptitude data on Air Force enlistees from 1956 through 1961, and reported that there had been a consistent upward trend in the overall aptitude level of enlisted accessions during that period. The data suggested that the selective recruiting program instituted in 1958 had in fact upgraded the aptitude quality of Air Force enlistees as compared with the previous enlistment base. It was further noted that, to the extent possible, the Air Force stood to gain qualitatively by enlisting only high school graduates. However, even during this period, it had not been possible to completely exclude all low aptitude enlistees, and duirng 1961 alone over 2,000 entered the Air Force.

Gordon and Flyer (1962) examined the performance characteristics of a group of over 11,000 low aptitude airmen who entered the Air Force during the first six months of 1956 and who had either successfully completed their four-year enlistment or had been discharged for unsuitability or nonadvancement. The objective of the study was to devise a brief screening battery that could be used to predict the success of such low ability recruits in their first enlistment, since it was recognized that the Air Force would continue to be faced with the periodic necessity of lowering its enlistment standards to meet its ever changing manpower quotas. It was found that a brief composite of aptitude tests and preservice educational level differentiated the successes from the failures quite well. It was also pointed out that the data supported a general policy of curtailing enlistment of low aptitude men, since almost 50% of them could not be expected to perform effectively over a four-year enlistment (a failure rate about twice that experienced for higher aptitude enlistees).

Later, McReynolds (1964) conducted a similar study of failure in basic training and found the rate of elimination for high school graduates to be only about one-fourth of that for nongraduates. Again, it was suggested that some additional screening should be employed to predict the potential success of nongraduates when Air Force quotas could not be met by enlistment of recruits who had finished high school. Vitola, Valentine, and Tupes (1967) conducted a new analysis summarizing the trends in aptitude

and educational data for Air Force enlistees, this time covering the period from 1962 through 1965. They found an overall upward trend in aptitude and educational levels of airmen accessions during this period, similar to that previously reported in Lecznar's analysis of the 1956 through 1961 accessions data. There were also expected indications of positive relationships between education level and measured aptitude, and between region of enlistment and aptitude. High school nongraduates were again found to have lower average aptitudes than enlistees who had completed high school or high school plus some college.

From these personnel studies carried out through the mid-sixties, it is apparent that the Air Force was basically following a policy of selectively recruiting its way out of the literacy problem by taking in only limited numbers of low ability personnel, and stressing the sustained enlistment of high school graduates. However, as will be mentioned later in this historical summary, the situation was about to change with the advent of Project 100,000 in 1966, and this change would bring about a renewed interest in literacy and technical writing problems within the Air Force research program on Job Performance Aids which has spanned the period from the middle 1950s to the present time, and which has comprised a major portion of the Air Force effort in the technical writing and useability of materials areas over the years.

RESEARCH & DEVELOPMENT ON JOB PERFORMANCE AIDS & RELATED TOPICS

The Air Force R&D program on Job Performance Aids (JPAs) originated in the middle fifties within AFPTRC, and it has been actively pursued ever since by a number of subsequent research organizations. Even with the various fluctuations in funding and available resources that have occurred over the years, the total scope of the program has been so extensive that the area can only be highlighted here. However, major aspects of the JPA effort have dealt with matters directly relevant to Air Force literacy and technical writing problems, and much of the initial interest in informational JPAs grew out of early concerns in areas related to the redesign of manuals, checklists, and other written job materials to improve their useability and intelligibility. You will recall that by the mid-fifties, there was an increasing Air Force awareness that many training and job materials were poorly designed and unnecessarily difficult. This problem was of particular concern in the maintenance area, where technical orders and procedural manuals were critical to successful job performance and proper repair and troubleshooting of complex equipment. The JPA concept was conceived as a potential solution to the problem. Essentially, JPAs were defined as special items or devices designed to enhance or guide on-the-job performance. These included a variety of instructional manuals, proceduralized checklists, detailed diagrams, or other information storage devices and materials which could be provided to the technician to maximize his capability to retrieve and use needed job information in the exact sequence required.

Berkshire (1954) developed and evaluated an early set of simplified troubleshooting materials for radar mechanics. These materials included a set of color-coded schematics of the equipment, along with detailed written troubleshooting instructions. Tryouts of the materials demonstrated that use of the troubleshooting aid produced considerable time savings and greatly reduced errors made by both experienced and inexperienced technicians. Subsequent work resulted in a series of reports providing guidelines for the improved preparation of maintenance information, job instructions, checklists, drawings, and diagrams (Folley & Altman, 1956) and maintenance manuals and handbooks (Miller, 1956 and Newman, 1957a, b).

By 1958, a substantial amount of work had been accomplished, including development of improved task analysis techniques for designing performance aids (Wulff & Newman, 1956; Gunn, 1956; and Emeson & Wulff, 1957), research on the design of troubleshooting guides (Hoehn & Saltz, 1956; Hoehn & Wardell, 1957; and Hoehn & Aukes, 1958), and a study of audio JPAs (Lumsdaine & Hoehn, 1958). Hoehn and Lumsdaine (1958) summarized the work done by AFPTRC on job aids up to 1958, including the concepts that had been developed concerning the role of handbooks, films, and other forms of aids in enhancing job performance and the techniques developed for systematically planning, preparing, and testing integrated job aids, training materials, and troubleshooting guides.

With the closing down of AFPTRC early in 1958, responsibility for the JPA work was moved to the training and engineering psychology research activities at Wright-Patterson AFB, and the continuity of the program was maintained in spite of a variety of organizational changes that occurred there through the late fifties. Effectiveness and useability of checklists (Rees & Kama, 1959) and technical manuals (Ross, D.A., 1959) were further evaluated, and research on improving the design and formatting of job guides and maintenance checklists was continued (Rees & Copeland, 1959 and Rees, 1959). By the early sixties a systematic JPA R&D program was again underway.

Folley (1961a) described the potential contribution of performance aids to improving system effectiveness and outlined a variety of research problems associated with integrating the design of such aids into the system development process. Folley and Munger (1961) reviewed the literature on the design of informational JPAs and noted that most of the prior work had involved development and tryout of various sample aids, as well as research on design and use of procedural aids for improving maintenance troubleshooting performance. Folley (1961b) presented a systematic procedure for developing performance aids which incorporated four key steps: identification of the task elements the aid should address, determination of required functional characteristics of the aid for each task element, specification of the physical design of the aid to perform the required functions, and evaluation for modifying or updating the aid as necessary.

The procedure was successfully applied to development of a set of aids for a missile system (Folley & Shettel, 1962) and was refined by further work through the mid-sixties.

During this period and on through the late sixties, research on the development and effectiveness of JPAs was pursued in order to expand the data base on their potential application and payoff to the Air Force. Elliott (1965) studied the effects of varying the format and level of detail of performance aids on a troubleshooting task and found that performance was generally better with a block diagram format than a list structure format, and that speed of task accomplishment was increased by use of a lower level of detail. Other studies demonstrated that proceduralized troubleshooting could produce acceptable or better performance on complex maintenance tasks than standard approaches, while also permitting time and training savings (Elliott, 1967). Folley and Elliott (1967) conducted a field survey of electronic maintenance technical data and concluded that the performance aid requirements of electronic techniques were not being adequately met, even though a variety of proven performance aid techniques and concepts were then available. A series of studies undertaken to provide an experimental comparison between proceduralized and conventional electronic troubleshooting (Elliott & Joyce, 1968) further demonstrated that the use of the proceduralized performance aid approach produced substantial savings in training time with no apparent decrement in task performance.

By the 1970s, the JPA approach had been sufficiently demonstrated that increased effort could be devoted to its dissemination and application, and various guidelines and handbooks on design and use of fully proceduralized JPAs were produced. These included the following: specifications for preparation of fully proceduralized JPAs (Chenzoff, Mallory, & Joyce, 1971; Folley, Joyce, Mallory, & Thomas, 1971a; and Joyce, Chenzoff, Mulligan, & Mallory, 1973a); handbooks for development of JPAs (Folley, Joyce, Mallory, & Thomas, 1971b; Foley, 1972; and Joyce, Chenzoff, Mulligan, & Mallory, 1973b); and handbooks for managing the procurement of JPAs in accordance with the specifications (Joyce, Folley, & Elliott, 1971 and Joyce, Chenzoff, Mulligan, & Mallory, 1973c). The JPA R&D program has continued through the present time and recent studies have attempted to develop improved JPA formats (Joyce & Chenzoff, 1974) and to demonstrate expanded applications of the JPA concept (Mullen & Joyce, 1974). Current efforts involve the development and evaluation of integrated sets of fully proceduralized JPAs for entire systems, such as the UH1H Helicopter and the C141 aircraft (Shriver & Foley, 1975; Foley, 1975a, b; and Shriver, 1975).

The techniques and concepts that have emerged from the JPA R&D program have had a positive impact on literacy and technical writing problems in the Air Force, and present programs to improve Technical Orders (TOs) and maintenance publications are applying principles growing out of JPA research. Also, the JPA approach appears to have produced a viable alternative for reducing the literacy demands of manuals and job materials which should see increasing application in the future.

OTHER RESEARCH ON READING AND READABILITY

A number of Air Force studies of reading and readability have occurred over the years that were conducted outside the ongoing research programs of the primary personnel, training, and human resources R&D organizations. One of the most significant among these was Taylor's (1957) development and validation of the "cloze" technique for measuring readability, which has seen widespread application in a variety of settings both within and outside the Air Force. Also notable has been the work of Davis at the Air Force Institute of Technology (AFIT) on the influence of a variety of textual, content, format, and personal variables on the effectiveness of technical communication via written materials (Davis, 1965, 1966, 1974a, b, c) and an AFIT thesis analyzing the relationship between readability of Air Force procedural manuals, literacy gaps, and frequency of discrepancies involving actual noncompliance with procedures on the job (Johnson, Relova, & Stafford, 1972).

RECENT RESEARCH ON LITERACY AND READABILITY

From the mid-sixties to the present time, the Air Force has maintained a very active R&D program in the literacy and readability area that has been carried on by both training and personnel researchers. A bit of background may help to place this recent research into some perspective. You will recall that from the mid-fifties through the mid-sixties the Air Force had been pursuing a selective enlistment program which had allowed recruitment on the basis of both aptitude and educational level and which had kept the overall quality of accessions fairly high. While limited numbers of low ability men had been enlisted throughout this period, they had not entered in sufficient numbers to present a major literacy problem. However, in 1966 the situation was radically altered with the initiation of Project 100,000 by the Department of Defense, and this brought increasing numbers of low ability personnel into the Air Force and other Services. The objectives of the program required that certain quotas of marginal personnel be accepted and further specified that these personnel be given the opportunity to enter a wide variety of career areas and job specialties that had not previously been open to them. Thus, Project 100,000 resulted in a rapid increase in input of men with marginal literacy and aptitude skills, and the Air Force was faced with the necessity of considering a variety of measures to adjust training and job demands to accommodate them. As a result, increased interest in research on literacy and readability problems emerged as a major concern.

Two problems became immediately apparent. First, there was little data on the reading abilities of Air Force personnel across career fields, and large scale reading testing on a long-term basis would prove costly for making assessments of the scope of the Air Force's emerging literacy problem. Furthermore, it would be difficult to do much about the problem of adjusting reading demands without some indication of the reading abilities of personnel, which could be used for recommending appropriate

readability levels for materials. Second, it would be similarly difficult to check the readability of the large amounts of written training and job materials produced by the Air Force without some automated alternative to the hand-calculated metrics that had been in use since the early fifties. Research had been initiated prior to the advent of Project 100,000 that provided timely answers to these problems.

Madden and Tupes (1966) had already developed conversion tables for estimating reading achievement from the AQE General Aptitude Index in terms of reading grade level as measured by the California Achievement Test and scaled score as measured by the Davis Reading Test. They had also provided distributions of estimated reading grade level for non-prior-service airmen and for airmen in 29 Air Force career fields. The conversion procedure for estimating reading ability has been in use ever since, and the career field data served as the standard source of target levels for readability of Air Force materials for almost a decade.

Also by 1966, an Automated Readability Index (ARI) had been developed (Smith & Senter, 1967 and Smith & Kincaid, 1970). The ARI provided an easy, automated method of collecting the data required to estimate readability of textual material by mechanical tabulation of word and sentence length as the material was typed on a standard typewriter. Inserting these data into a simple formula provided an index of the reading difficulty of the material. The feasibility of using the ARI to assess the comprehensibility of Air Force technical orders was subsequently demonstrated (Kincaid, Yasutake, & Geiselhart, 1967) and the ARI has been periodically used for that purpose and various other applications up to the present time.

In 1968, the Air Force Human Resources Laboratory (AFHRL) was established, and the Technical Training Division of the Laboratory was given prime responsibility for Air Force reading, literacy, and readability research. The Division has been actively pursuing work in these areas since its establishment in 1969. From its inception, this R&D program has been planned and carried out with a long-term view toward researching the literacy problem systematically; that is, to have the research address major aspects of the problem from both a men and a materials standpoint.

Initial effort was placed on researching techniques for improving materials and readability. Sellman (1970, 1972) evaluated a modified career development course (CDC) format to determine if reducing the reading difficulty of the material, including more illustrations, and providing audio supplementation of the text would improve the CDC as a training device for lower ability airmen. The simplified materials with audio supplementation produced significanlty better learning scores for high, middle, and low aptitude groups, with the greatest gains obtained by the lower ability trainees. At about this same time, Huff and Smith (1970) compared the reliability of the Fog Count with that of the Automated Readability Index (ARI) and established ARI baseline data on a sample of

Air Force CDCs. The Fog Count, then the accepted Air Force readability measure, was found to be very unreliable with coefficients ranging from only .49 to .56. Reliability coefficients for the ARI, on the other hand, ranged from .98 to .99+, and it was concluded that the ARI offered an extremely reliable tool for assessing the relative readability of materials. It was also pointed out that the ease and rapidity with which required readability data could be collected was another advantage of the ARI, and a set of instructions for its use was provided.

Additional emphasis was then placed on readability research. An extensive study of available readability measures was undertaken (Williams, Siegel, & Burkett, 1974) and a report was prepared summarizing a variety of techniques for use in evaluating and improving readability of materials (Siegel, Federman, & Burkett, 1974). This work served as the foundation for a subsequent series of efforts, still underway, to develop improved measures of textual comprehensibility based upon analysis of key cognitive and psycholinguistic factors that contribute to difficulty of written materials (Siegel, Lambert, & Burkett, 1974 and Siegel & Burkett, 1974). Eight cognitive/intellective factors and six psycholinguistic factors were identified that could serve as the basis for a new approach to measuring comprehensibility. Measures of these factors were derived and experimentally validated, and current work is attempting to further refine these metrics with a view toward ultimate development of a set of computerized comprehensibility measures.

The early interests in materials and readability also generated a somewhat different line of research into ways of reducing reading demands through use of specially designed media, simplification of materials, and audio/visual techniques. A systematized audio/visual approach to self-paced job training called Automated Apprenticeship Training (AAT), which employed an easily operated portable sound/filmstrip teaching device, was developed (Pieper, Catrow, Swezey, & Smith, 1973). The AAT format was specifically designed to reduce dependence on reading skills, and it was shown to be highly effective for low aptitude trainees without penalizing the higher ability men going through the same training program. The AAT approach showed good potential as a means of reducing the reading demands placed on marginal airmen, and its self-paced audio/visual features also produced substantial training time savings, with performance equal to or better than conventional approaches. In another attempt to systematically reduce reading demands, a series of experiments was conducted to examine the effects of lowering the reading grade level of textual material and providing tape recordings of the text for increasing comprehensibility (Siegel, Lautman, & Burkett, 1974). Greatest benefit was found for just reducing the reading grade level difficulty of the written material, and this was more effective for self-study materials than for resident course study guides. This study demonstrated the value of matching the difficulty of materials to the reading ability of the men, but did not show any particular advantage for additional audio supplementation of the material.

By 1973, there was an increasing Air Force indication that the reading problems of some airmen were causing problems in both resident training and on-the-job training. Accordingly, studies of the reading abilities of airmen and the scope of the reading training problem were undertaken. Mockovak (1974a) surveyed Air Force reading improvement programs and found that from 1 April 72 to 1 April 73, 5744 airmen participated in such programs, with the most frequently cited problems involving inability to read and pass CDCs. He also found that ten career fields accounted for over 80% of the problem readers who had required the training. These included Aircraft Accessory Maintenance (42XXX) 6.2%, Aircraft Maintenance (43XXX) 18.7%, Mechanical/Electrical Trades (54XXX) 3.9%, Structural/Pavements Trades (55XXX) 8.0%, Transportation (60XXX) 11.2%, Food Service (62XXX) 5.9%, Fuel Services (63XXX) 5.1%, Supply (64XXX) 8.3%, Administration (70XXX) 12.7%, and Security Police (81XXX) 3.9%. In a related study, Mockovak (1974b) demonstrated a methodology for determining the reading skills and requirements of Air Force career fields, and for identifying areas where literacy gaps existed. Basically, the methodology involved conducting a readability analysis of a sample of the reading ability of men in the career field using the Madden-Tupes conversion from aptitude scores procedure previously mentioned in this review. By computing differences between the readability of the materials and the reading ability of the men, literacy gaps could be determined and reading requirement levels derived.

At about this same time, the reading skills acquisition process was examined through research on a developmental model of auding and reading skills (Sticht, Beck, Hauke, Kleiman, & James, 1974). The model suggested that reading was based upon and utilized the same conceptual base and languaging skills as were used in auding (i.e., comprehending spoken language) and that ability to comprehend by reading could be improved by improving the conceptual base and language skill by means of auding. Implications were also found for literacy diagnostic testing, in that it was apparent that some reading problems may really be the result of poor language skills. It was thought that the latter might be indexed by means of an auding test. Then, differences between a reading and language skill, indexed by the auding test, might be useful in prescribing reading or language training for individual airmen. A follow-on effort is now underway to develop such tests.[1]

The personnel and materials areas were later integrated into an extensive effort designed to determine the reading requirements across Air Force career fields and to identify problem areas where literacy gaps existed. Mockovak (1974c), using the methodology he had developed earlier, compiled reading and reading requirement data for 56 career ladders and reading ability target data for all 277 career ladders in the Air Force.

---

[1]See the paper by Groff in these proceedings.

Across the 56 ladders, average reading requirement level (RRL) obtained was 12.3 reading grade level, ranging from a low RRL of 10.6 for cooks to a high RRL of 14.0 for telephone switching equipment repairmen. Average reading ability of personnel across all 277 ladders was 10.8, but there was considerable variation between ladders, with some ranging as low as 8.5 and others as high as 14.5. However, meaningful target levels for writers were obtained, and these are currently being used in the preparation of Air Force training and job materials, replacing the older Madden-Tupes data that had been in use since the mid-sixties. An average literacy gap of around two reading grade levels was also found for the ten problem career fields previously noted, suggesting that such mismatches were, in fact, contributing to the Air Force's reading problems.

Work is currently being conducted on both readability of materials and on reading skills of personnel. Improving readability is the goal of present attempts to develop new textual comprehensibility measures (Siegel & Wolf, 1975). Reading skills of personnel are being addressed through development of literacy assessment tests for improved diagnosis of individual reading and auding problems and by efforts to develop and demonstrate prototype job-oriented reading training programs.

It is anticipated that future research in the AFHRL program will continue to follow a "systems approach". That is, it is felt that if headway is to be made, systematic research on major aspects of the literacy problem from both a personnel and a materials viewpoint should be pursued. In dealing with the personnel, research is anticipated to address the problems of identifying better ways to measure their literacy skills and of developing training techniques designed to upgrade their reading skills in a job-meaningful manner. Similarly, in the materials area, major thrusts are anticipated to involve both research on measurement of the comprehensibility and reading demands of job materials and on development of improved methods for preparing and modifying materials and media to appropriately useable literacy levels.

CONCLUSION

The phase-out of Project 100,000 and recent increases in enlistment standards have eased the literacy problem considerably, but recent staff studies and research findings indicate that the Air Force still has literacy and technical writing problems of sufficient magnitude to be of some concern. Present indications are that the advent of an all-volunteer force has neither resolved the problem nor made it particularly worse (Vitola, Mullins, & Brokaw, 1974) as some had originally supposed it might. However, a major policy change or mobilization such as those that have occurred in the past could again rapidly intensify the problem. While the prior R&D accomplished by the Air Force on literacy, technical writing, and related matters has provided a useful technology base and a variety of useful applications in these areas, no panaceas to the problems have been found. Gaps in our knowledge and techniques still exist which should

be addressed to provide the fools required to deal with problems in these areas, and future Air Force literacy and technical writing R&D will be oriented toward these technology gaps and toward those specific priorities and requirements for applied research in these areas that emerge in the years to come.

COMMENTS ON THE PAPER BY BURKETT

John Guthrie
International Reading Association

Guthrie first reviewed the highlights of Burkett's paper, beginning with Burkett's statement that the basic problem which has been the focal point for much of the Air Force's literacy and readability research is a literacy gap between the reading ability of personnel and the reading difficulty of materials. Burkett then reviewed work in readability of materials and literacy training in the 1950s, pointing out that neither research on readability nor on literacy training had indicated major impact on people's ability to perform their jobs. A second wave of studies concerned personnel factors related to job performance and showed that high school graduates had higher success rates than non-high school graduates, leading the Air Force to recruit high school graduates. However, Project 100,000 produced an influx of lower aptitude personnel, and once again a proliferation of research on how to best train and utilize such personnel occurred. This work included the development of job performance aids to reduce cognitive and literacy demands of jobs and job training programs. Other work, including the current work, focuses on the one hand on how to circumvent the person's literacy problem by reducing the reading demands of jobs or improving the readability of material and on the other hand on the assessment and improvement of the literacy skills of personnel through improved literacy training. In this regard, Guthrie considered the Air Force's and other Armed Services' literacy problems as analogous to that faced by some developing countries. In such countries there is a literacy gap: a mismatch between the capabilities of the people and the demands of a modern agricultural or technological society. It is a goal of underdeveloped countries to develop their economic and social stature, and literacy training is viewed as a means to that end, rather than an end in itself. Thus, to achieve social stature there may be a need to increase agricultural productivity. In order to increase agricultural productivity, there is a need for farmers who can read instructions for new fertilizers, who can calculate the properties of mixtures, and who know enough to know when and when not to apply the fertilizers. For instance, in one agricultural country, cotton is a cash crop; therefore, a major literacy program focuses on teaching people to read within the context of cotton growing. Another teaches reading within the context of banana growing, and so on. Without overdrawing the parallel between the Armed Services and the developing countries, Guthrie suggested that in both cases there is a job to be performed — cotton growing or troubleshooting a radar set — as the primary interest. Also, performance of the job requires job knowledge in addition to relevant cognitive skills, such as reading, and motivation to perform the job. Thus, job performance requires job knowledge, reading, and motivation. Hence, training programs may require the teaching of all three of these factors in some integrated form.

ONGOING RESEARCH AND DEVELOPMENT
IN LITERACY TRAINING
IN THE ARMED SERVICES

INTRODUCTION

The papers of the preceding section provide a perspective on past R&D efforts on reading and on the ways in which literacy training has been conducted within the Services. The present section summarizes the current situation regarding literacy training and ongoing R&D projects to understand and manage problems of literacy within the Services.

In his review of the Navy's ongoing program of literacy research, Tom Duffy discusses two major efforts: those aimed at providing a data base to characterize the literacy problem in the Navy, and those which examine the Navy's reading training programs and the people in them. In the first of these two efforts, the research investigated reading ability and its relationship to personnel characteristics (education, race, and scores on the ability tests given in the Navy Basic Battery Test) and to the readability of Navy manuals. Related studies looked at the relationship between reading ability and two measures used in the Navy to measure job effectiveness: attrition from recruit training and performance in "A" School.

In presenting the second effort, Duffy describes three current Navy and Marine Corps literacy training programs and then discusses the meaning of student reading grade level gains in terms of entry selection, racial and educational characteristics of the students, characteristics of the programs, gains in civilian reading programs, and the criterion and ultimate goals of literacy training programs. Finally, he briefly describes three experimental programs and research efforts to characterize the successful student in the Navy program in San Diego.

The research and development program in the Army parallels that of the Navy in that one phase has dealt with the nature of literacy problems and a second phase has dealt with a literacy training program. John Caylor summarizes the Army program, which consisted of (1) an earlier, primarily research phase that sought to define the literacy problem in terms of actual job functional literacy demands, and (2) a later, primarily developmental phase that produced a job functional literacy training program aimed at meeting those demands.

The literacy problem has two sides, as Burkett has stated in his historical review of Air Force R&D. The Army program has also looked at both sides of that problem. To determine the job functional reading requirements of selected Army jobs, the research studied the relationship of general reading ability to measures of job proficiency, estimates of the readability of Army manuals in the selected job areas, and performance on job reading task tests. These three measures were in general agreement with each other. The reading ability of Army personnel in Mental Categories III and IV was assessed directly by a standardized reading test. A correlation study showed that reading ability could also be assessed indirectly from AFQT scores. The results of the research phase showed that the reading demands of Army jobs far exceeded the reading ability of many Army personnel.

Preceding page blank

The information on the nature of the literacy problem was used to develop a functional literacy program (FLIT) which provides job reading training using the content and reading materials of the student's job area. One segment of the program stresses the application of the student's existing general reading skills to specific kinds of job reading tasks, using actual job reading materials. Another segment uses specially prepared job-oriented reading passages and exercises to improve basic reading skills. This program was implemented at all of the Army Training Centers.

Steve Groff opens his survey of ongoing literacy research and development in the Air Force by looking at the current Air Force literacy training. During basic military training, reading training is provided in Reading Proficiency Units, which offer general literacy training. After assignment to a permanent duty station, reading training may also be obtained through the base education office. These reading programs vary greatly. The high numbers of people enrolled in literacy training and the results of studies which have found a literacy gap between the estimates of job reading demands and the reading ability of the personnel in those jobs all point to a definite literacy problem in the Air Force, despite high enlistment standards.

Current Air Force R&D involves two major efforts. The first of these is the development of a Literacy Assessment Battery which can be used to identify people who would most likely benefit from further literacy training. The Battery yields an indication of the discrepancy, if any, between the person's auding (oracy) and reading (literacy) skills. The person whose oracy skills exceed his literacy skills is the most likely candidate for literacy training. Such a person has a greater command of language skills than his reading score alone would suggest, and literacy training could bring his literacy skills up to the level of his oracy skills. The second ongoing effort is a prototype job-oriented reading program for the Air Force. This program is based on the approach used in the Army FLIT program, but it focuses on the higher level comprehension skills.

Groff points out that no one literacy program meets all literacy needs for all people. A total career education program would include literacy training, not as a one-time shot, but available throughout an airman's career, in different forms for different needs. Such a program could include the Reading Proficiency Unit training to improve basic reading skills, the job-oriented reading program to improve conceptual behavior in a job context, and general reading/GED programs to provide access to higher levels of skill.

LITERACY RESEARCH IN THE NAVY

Thomas M. Duffy

Navy Personnel Research and Development Center

Literacy research in the Navy today has had two major objectives. First, there have been efforts to characterize the literacy "problem" in the Navy and thereby provide a data base for evaluating alternative approaches to insuring that a literacy deficit at any level does not hamper the effectiveness of the service. The second major line of research has focused on the Navy's current approach to deficient literacy skills — reading training programs for personnel with marginal reading ability. Each of these avenues of research are discussed below.

CHARACTERIZING THE LITERACY PROBLEM IN THE NAVY

In this research, a group of us at NPRDC have initiated efforts to characterize the reading skills of personnel and examine basic demographic data on reading and its relationship to other abilities and background characteristics. This research has also involved an examination of the relationship of reading skill to job performance.

Interrelationships Among Reading and Other Personnel Characteristics

In a joint effort with the Naval Training Center and Recruit Training Command in San Diego, we have now collected reading and other ability and background data, as well as performance data, on some 25,000 recruits and 1500 men receiving occupational ("A" School) training.

For both the recruit and A School samples, we obtained all of the test and background data collected as a normal part of processing into the service. These data include the following test scores and background information:

1. General Classification Test (GCT) — A test of general ability involving verbal analogy and sentence completion items.
2. Arithmetic Reasoning Test (ARI) — A test of mathematical ability involving word problems.
3. Mechanical Ability (MECH) — A low-verbal test of knowledge of mechanical principles.
4. Clerical Test (CLER) — A speeded digit search test requiring no verbal skill.
5. Electronics Test (ETST) — A verbal test of electronics aptitude.

6. Shop Practices Test (SHOP) — A test of knowledge requiring the matching of a picture of a tool to verbal descriptions of uses.

7. Armed Forces Qualification Test (AFQT) — A score derived from the GCT, ARI, MECH, & CLER by their regression on the former AFQT test of general ability which was administered throughout the Armed Forces.

8. Years of Education — Self reported.

9. Race — This data was obtained only for the recruit sample.

In addition to the routine personnel data, all of the A School sample and 1,294 recruits were also tested on the Navy Pattern Matching (PM) test of nonverbal ability. The PM was derived from the Raven's Standard Progressive Matrices Test (Raven, 1958) which is considered by some to be a culture-free test of nonverbal intellectual ability (Jensen, 1972; Carver, 1973d; Spearman, 1946).

Recruits' reading grade level scores were obtained using the vocabulary and comprehension subtests of the Gates-MacGinitie Reading Test, Survey D (Gates & MacGinitie, 1965). The A School personnel were tested using the vocabulary and comprehension subtests of the Nelson-Denny Reading Test (Nelson & Denny, 1960).

Reading Levels

Figure 1 presents the distribution of reading levels in the recruit sample. The skewness of the distribution reflects the limitation of the Gates-MacGinitie test (which has a maximum score of 12.0 RGL) for this sample.

The median RGL is 10.7, which is generally reflective of the education level in our sample, in which 70.9% left school after either the 11th or 12th grade. However, the distribution of scores indicates that 18% of our sample has less than an 8th-grade reading skill. In order to get some idea of the implications of these reading scores, we did a readability analysis of those manuals and tests which most recruits are expected to read. The materials were analyzed using the FORCAST readability formula (Caylor, Sticht, Fox, & Ford, 1973). The materials assessed and their reading difficulty were as follows: the Bluejackets' Manual (11.5 RGL), which is the basic manual in boot camp; the Airman (10.5), Fireman (10.2), and Seaman (10.2) rate training manuals used during and right after completing boot camp; and the general classification test (10.8 RGL) and arithmetic reasoning inventory (9.0 RGL) which are administered in the process of classifying recruits.

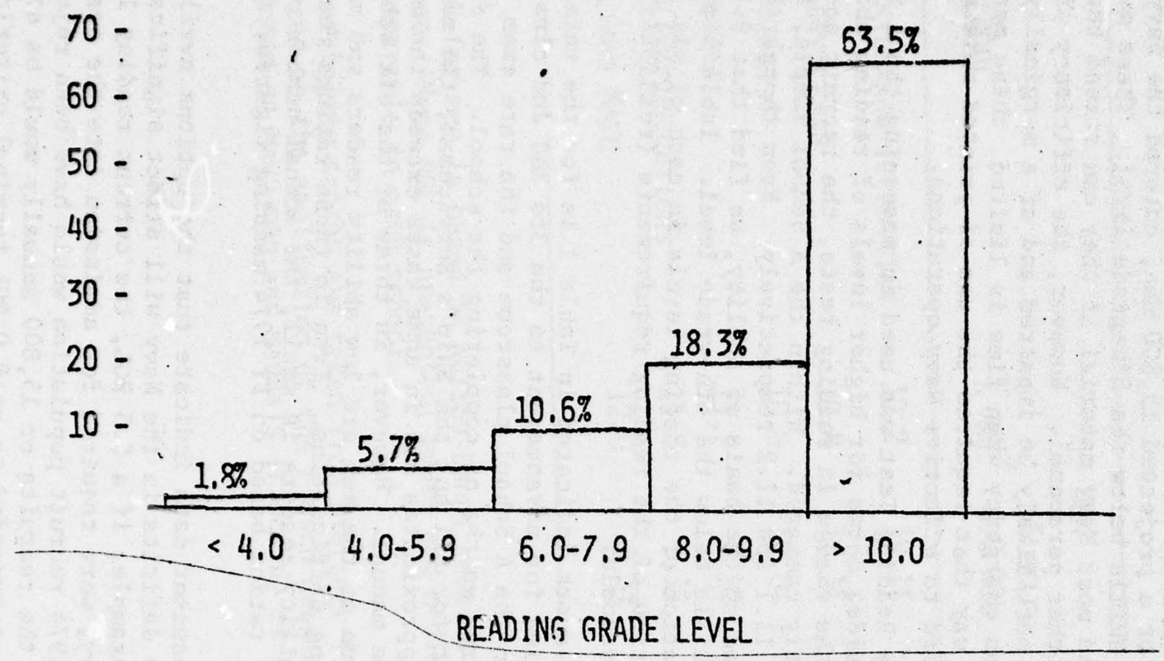


Figure 1. Distribution of Reading Levels of Recruits at San Diego Recruit Training Command (n = 24,729).

A comparison of the reading and readability levels indicates that 40% to 50% of the recruits have a reading ability less than that demanded by their job reading material. While the effects of this apparent "literacy gap" on comprehension and performance is uncertain, it seems likely that the 1.8% of recruits (a projected 1800 men for FY 1974) who read below the 4th-grade level will have great difficulty, if not total impossibility, in comprehending these job reading materials. Further, in our recruit sample, 18%, or a projected 15,800 men, entered the Navy during FY 1974 with reading skills below the 8th-grade level. These men may be expected to comprehend most Navy material if they can reread the material and are assisted by other personnel. However, the efficiency of their job performance will most likely be impaired and of a marginally acceptable level. If, in an emergency when time is limited, these men are required to act in any way that requires the use of printed material, they may well prove a hazard to effective Navy operations.

The Nelson-Denny Reading Test was used in assessing the A School sample, since it provides norms for higher levels of reading ability. However, because of the change in reading tests, the recruit and A School data cannot be directly compared. Within the A School sample, the mean and median RGLs were 11.1 and 11.2 respectively. Even though the A School personnel are selected on the basis of ability, we find that 9.6% of the sample had reading skills below the 8th-grade level. Table 1 presents a more detailed examination of the reading levels in each of six A Schools, as well as an indication of the reading requirements (readability of materials) in those schools.

The readability score indicated in Table 1 is for the rate training manual which is studied for advancement to the 3rd and 2nd class rates. The manual is used in the A School classroom and the rate exam is typically taken within three months of completing the school. The data in Table 1 indicate that for all but the Ship's Serviceman, the mean reading ability of the men approximates and in some cases exceeds the mean reading difficulty of the manual. However, in three of the six schools, a significant proportion of the men are low ability readers and may be expected to face reading difficulties. Even in those ratings where the proportion of low ability readers is small, the actual number of men, several hundreds per rating based on FY 1974 manning figures, is considerable.

All of our assessment data indicate that the options available in approaching literacy deficits in the Navy will affect significant numbers of personnel. For example, if a 5.5 RGL, the current reading level deemed necessary in the Navy, were required for admission into the service, 9% or 7,800 of the FY 1974 recruit population would have been rejected. Eighteen percent of the recruits or 15,800 annually would be eligible for a literacy program expanded to an 8.0 RGL terminal criterion, a criterion which only approaches the initial reading demands in the service. The same number of men would be affected by the implementation of limited

TABLE 1. READING GRADE LEVELS (RGL) OF MEN & READING GRADE LEVEL REQUIREMENTS FOR MANUALS IN EACH OF SIX NAVY OCCUPATIONAL TRAINING "A" SCHOOLS.

	MEN		MANUALS
SCHOOL	$\overline{X}$ RGL*	% BELOW 8.0 RGL	$\overline{X}$ RGL**
HULL MAINTENANCE TECHNICIAN	10.2	19.1%	10.7
SHIP'S SERVICEMAN	10.4	18.1%	12.9
MESS SPECIALIST	10.6	12.0%	10.2
QUARTERMASTER	11.7	2.9%	10.9
INTERIOR COMMUNICATIONS	11.9	5.1%	12.6
ELECTRICIAN'S MATE	12.0	2.3%	12.7

*RGL for men is based on Nelson-Denny reading test performance. Mean RGL is presented; means and medians were comparable in each school.

**Readability scores are taken from Biersner (1975) and are based on the application of the Flesch Reading Ease formula normed on Navy men and manuals (Kincaid, et al., 1975).

duty assignments for below-8th-grade readers. The costs involved and the number of men affected by implementation of any of these options demands that the options be carefully considered.

Our findings with regard to literacy skill levels in the Navy indicate that one or more of the above options must be implemented to assure effective performance levels in the service. These options, however, must necessarily focus on the very low literate man in order to be manageable and within reasonable cost figures. The distribution of reading skills in the A Schools suggests an additional option which would address personnel with less severe literacy deficits. This option involves a literacy training program for personnel deficient in reading skills, but otherwise qualified for A School. The program would have as a *criterion, the ability demands of each school and would provide functional reading training* preparatory to, or in conjunction with, school training. The major benefit of this program is that it would provide the opportunity of upward mobility and career flexibility for men who, while eligible for A School training, are likely to perform at a marginal level. Additionally, the program would involve fewer personnel and hence, would be more manageable than a broad-based literacy program. HumRRO, under contract to the Center, is currently evaluating the cost-effectiveness and management considerations involved in this and other literacy training options.

Education. It is frequently asserted that the current move toward accepting only high school graduates will greatly reduce or eliminate the marginally literate personnel, and hence, there will be no need for Navy commitment to literacy training. Our data indicate, to the contrary, that a high school diploma bears little relationship to reading skills. The median reading levels for our high school and non-high school graduates differ by less than one RGL (the medians are 10.9 and 10.2 respectively). Figure 2 presents the distribution of reading grade levels. While proportionately fewer high school graduates fall below an 8th-grade RGL, this proportion (16.9%) nonetheless indicates that the selection of only high school graduates would still produce a significant number of marginally skilled individuals.

Further analysis on a sample of 19,000 recruits indicates that years of education, ranging from 8th to 6th grade, correlates only 0.13 with reading ability. Of all of the test and background information for which we have data, years of education shows the least relationship to reading. This finding may be contrasted with previous research which has found years of education to be the most valid predictor of attrition from the service (Plag & Hardacore, 1964) and delinquency (Gunderson & Ballard, undated). One might conclude from these data that a high school diploma indicates a person's willingness to conform to the rules of society rather than his ability. However, a more likely explanation is that over recent years the predictive power of years of education has been reduced due to the increasing porportion of high school graduates entering the Navy. Evidence for this hypothesis comes from the additional finding in our

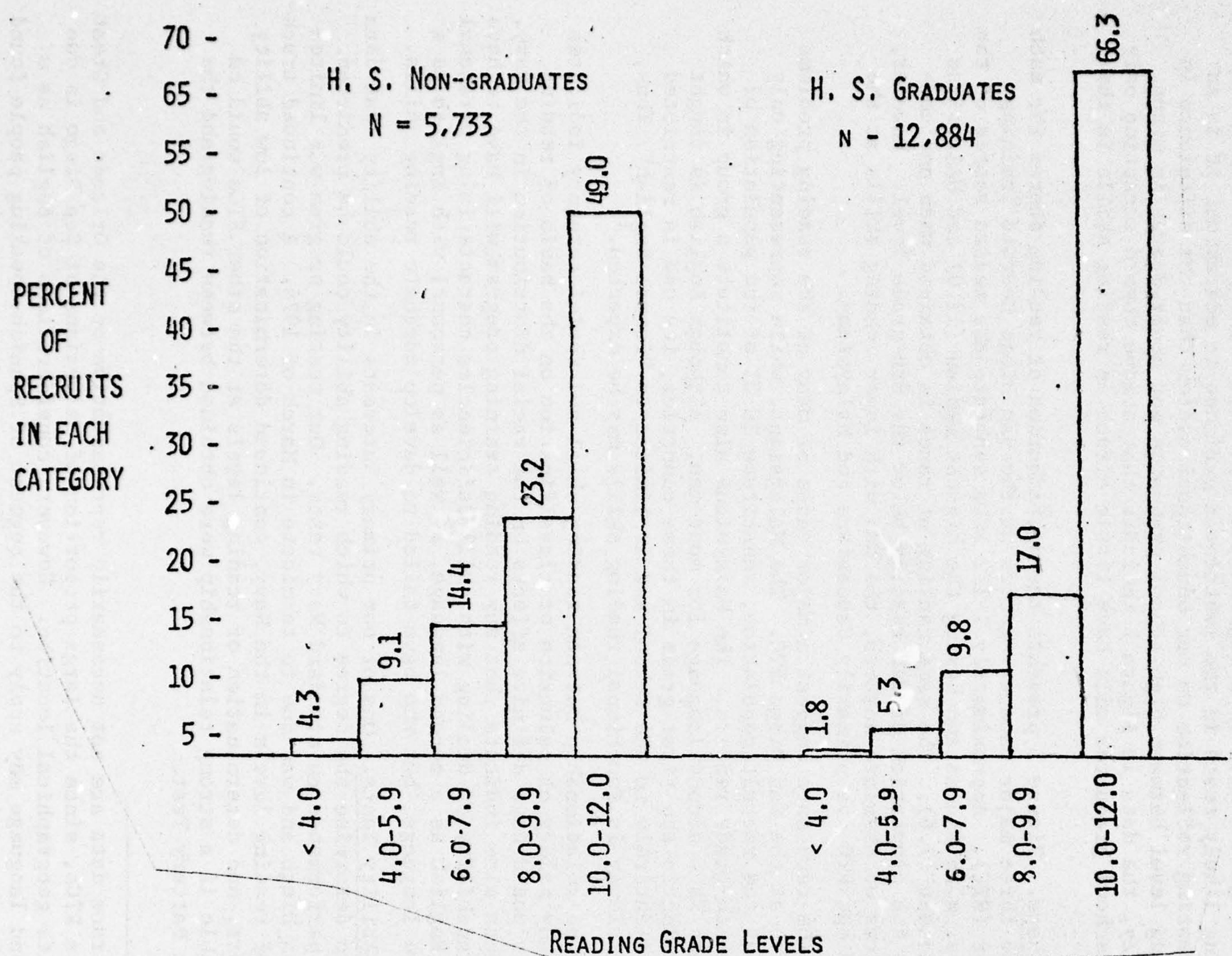


Figure 2. Reading Levels for High School Graduates & Non-Graduates at RTC San Diego: May-Oct 1974 (n = 18,617).

analyses that years of education is no longer a strong predictor of attrition. The reduced variance in years of education is indicated by the fact that 77% of our sample completed 11 to 13 years of education. Even though the lack of a significant correlation between education and reading likely rests in the restricted variance in education, it is an interesting reflection on our educational system that the difference in reading level between high school graduates and nongraduates is minor. Clearly, the data in Figure 2 indicate that a move toward accepting only high school graduates will have little effect on reading skills in the Navy.

Race. Figure 3 presents the distribution of reading scores for each of the three major race categories at the San Diego Recruit Training Center (RTC). Approximately 1-1/2 RGLs separate the median scores of the groups, with Caucasians having the highest median (11.0) and Malaysians the lowest (7.6). The same ranking of races is obtained when one compares the proportion of men reading below the 8th-grade level. However, in terms of absolute numbers, the men with lower reading skills at the San Diego RTC are primarily Caucasians and Malaysians.

The race data suggest a major cause of many of the reading problems for men at the San Diego RTC. The Malaysians, while representing only 8.6% of the recruit population, constitute 28.0% of the population of below-8th-grade readers. The Malaysians also constitute a group in which English is a second language for most men. Although English is taught beginning in the first grade in these countries, its use is restricted almost entirely to the school and is limited to "formal English". Thus, a deficiency in functional reading skills may be expected.[1]

The relationship between reading skill and racial category indicates that any policy of selection or classification on the basis of reading skills would have definite effects on the racial distribution in the Navy. The data also indicate that any reading training program will have to have the capability of dealing with skill deficiencies characterizing personnel with English as a second language, as well as personnel with English as a native language, but who have failed to develop adequate reading skills.

Ability Tests. One of our primary interests in the ability test data was to determine the degree to which reading ability could be predicted from performance on standard Navy tests. Our testing program was limited to San Diego and was due to terminate in March of 1974. A continued tracking of reading levels in the Navy, continued determination of low ability readers, and determination of reading levels at the other RTCs would be possible if a strong relationship were obtained between reading and the Basic Battery Test.

---

[1]The race data are not necessarily representative of the Orlando and Great Lakes RTCs, since the large proportion of Malaysians at San Diego is due to its geographical location. However, considerations of English as a second language may apply to the segment of Spanish-speaking people found at those training commands.

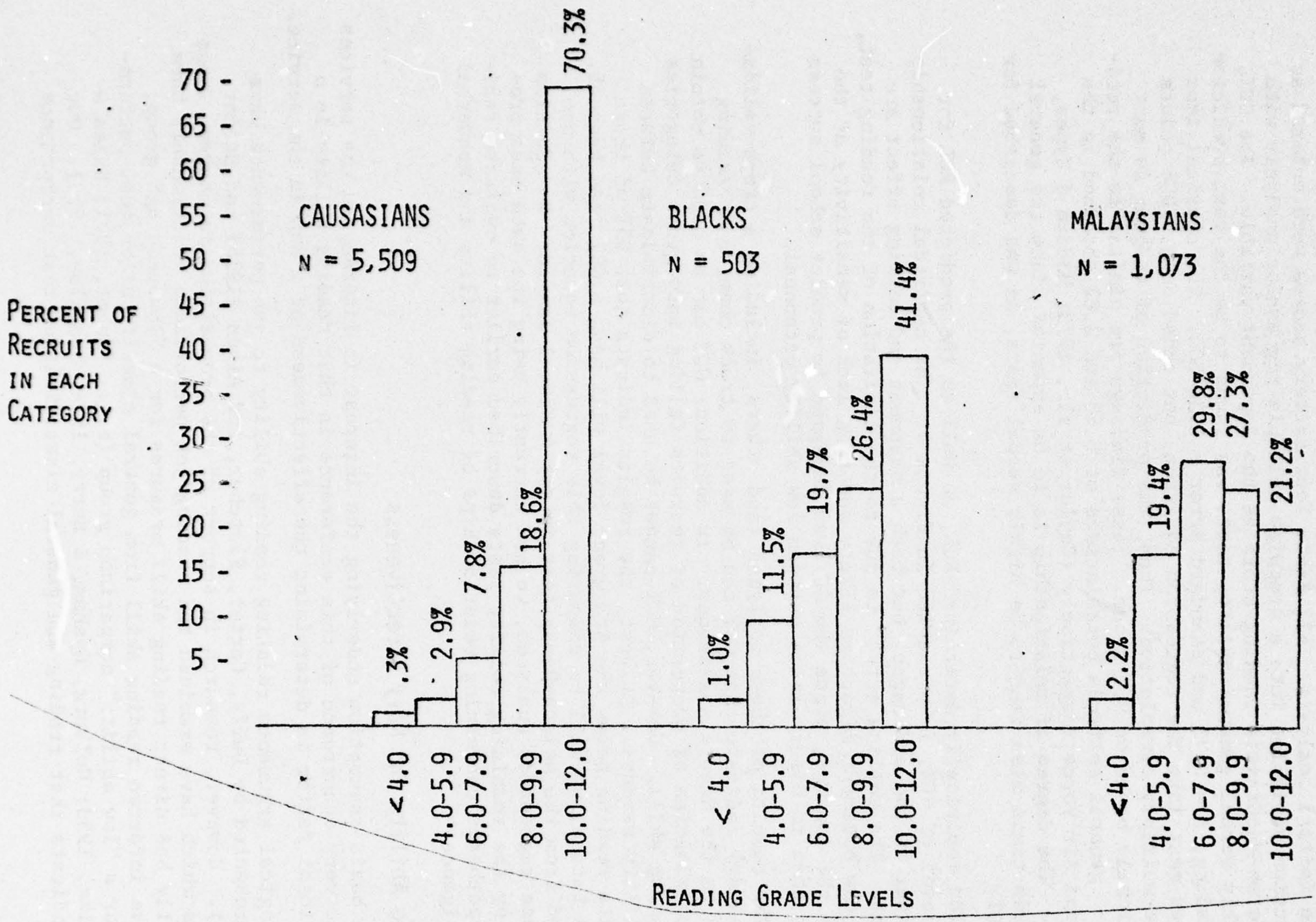


Figure 3. Reading Levels in Major Race Categories at RTC, San Diego: May-Aug 1974.

Data from the first 7093 recruits taking the reading test were used in the initial analysis. All Basic Test Battery scores were entered as prediction variables into a stepwise multiple regression analysis with the Gates-MacGinitie reading score as the dependent variable. The GCT, a highly verbal general ability test, was found to be the best predictor of reading, r = 0.73 and standard error = 1.36 RGL. The clerical test entered next into the regression equation, but added only 0.006 points to the multiple correlation. Thus, the prediction of reading is most efficiently based on GCT alone. These findings are similar to the reading — general aptitude correlations of 0.68 and 0.82 obtained in the Army and Air Force respectively (Caylor et al., 1973; Madden & Types, 1966). The degree of relationship is to be expected since the general aptitude tests also tend to be highly verbal tests, as was described for the GCT.

The empirically determined RGL, as well as the predicted RGL for each level of GCT, is presented in Figure 4. The empirical relationship is linear in the midranges, but both a basement and ceiling effect are evident. The ceiling effect is due to the limitation of the reading test, while the basement effect is likely due to a lack of sensitivity of the GCT at the low end, since the GCT was designed to predict school success rather than to discriminate between low ability personnel.

The results of these analyses (and others, including a cross-validation study) indicate that GCT can be used to track changes in reading skills in the Navy's assessions. In addition, GCT can be used to obtain an approximation of proportion of recruits falling into gross categories of reading skill. However, GCT cannot be used to discriminate between low ability readers. At best, the results indicate that 95% of those recruits reading below the 6th-grade level will have a GCT less than 45. In the future, we will be comparing this regression equation with one derived from the Nelson-Denny test in our A School sample. In addition to these regression analyses, we are currently using the data base provided by the remaining ability tests described earlier to evaluate various hypotheses concerning relationships of reading ability and nonverbal intelligence.

Reading Ability and Navy Effectiveness

A basic assumption underlying the interest in literacy in the services and the very occurrence of this conference is that reading ability is a significant factor in determining the effectiveness of a man in the service.

Logical arguments relating reading ability to job performance have been presented by Duffy, Carter, Fletcher, and Aiken (1975) and Sticht (1975a). However, research in support of the proposition is meager. Those studies which have examined the reading/job performance relationship have typically had direct reading skill measures for a "low reading" group, but have inferred reading skill from general classification test performance for a "low ability" comparison group (e.g., Fisher, 1971; Hagen & Thorndike, 1953; Hoiburg, Hysham, & Berry, 1974; Standlee, 1954). Our data indicate that reading and general classification test performance

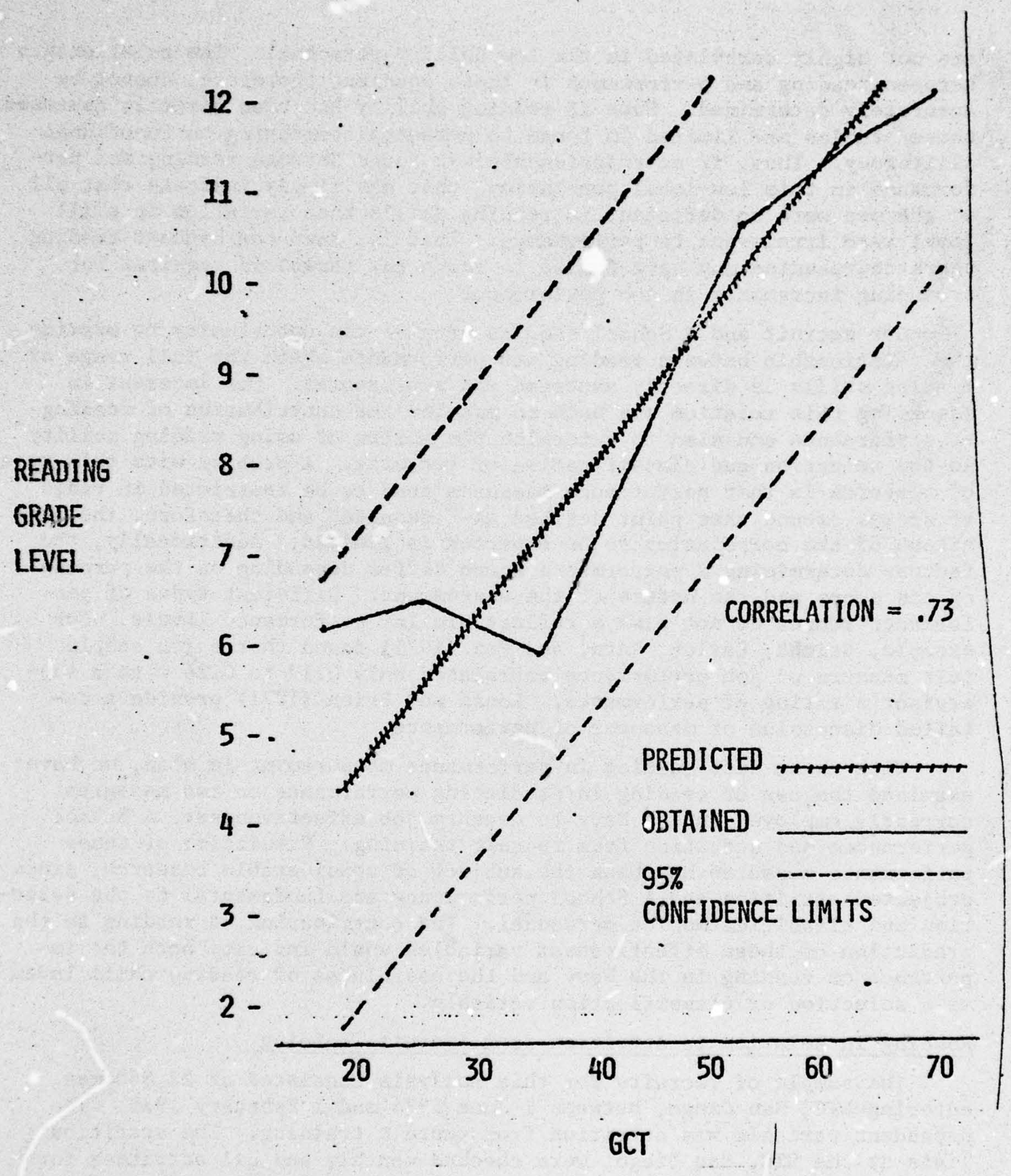


Figure 4. Relationship Between GCT & Reading Grade Level for 7093 Recruits.

are not highly correlated in the low ability personnel. The relationship between reading and performance in these studies, therefore, cannot be accurately determined. Even if reading ability had been directly assessed, these studies are limited in focus to personnel bordering on functional illiteracy. Thus, if no relationship was found between reading and performance in this low-level population, this may simply indicate that all of the men were so deficient in reading skills that variation in skill level were irrelevant to performance. That is, even the highest reading score represented may have failed to reach the threshold required for producing increments in job performance.

Our recruit and A School samples provide the opportunity to examine the relationship between reading and performance where the full range of reading skills is directly assessed and represented. Our interest in assessing this relation was both to examine the contribution of reading to performance and also to determine the merits of using reading ability in the selection and classification of recruits. A problem with this type of research is that performance measures tend to be restricted in range to scores around that point defined as "adequate" and therefore, the magnitude of the correlation to be expected is limited. Additionally, the factors determining a performance score differ depending on the purpose of the score and the nature of the assessment. Different types of performance scores do not always reflect similar performance levels. For example, Sticht, Caylor, Kern, and Fox (1971) found that a job sample test measure of job performance correlated only 0.13 to 0.24 with a supervisor's rating of performance. Ronan and Prien (1971) provide a detailed discussion of measures of performance.

With these difficulties in performance measurement in mind, we have examined the use of reading in predicting performance on two measures currently employed in the Navy to measure job effectiveness: A School performance and attrition from recruit training. Prediction of these performance measures has been the subject of considerable research, since projected attrition and A School performance are fundamental to the selection and classification of personnel. The contribution of reading to the prediction of these effectiveness variables would indicate both the importance of reading in the Navy and the usefulness of reading skill level as a selection or classification variable.

<u>Reading in Relation to Attrition from Recruit Training</u>

The sample of recruits for this analysis consisted of 22,840 men entering RTC, San Diego, between 1 June 1974 and 1 February 1975. The dependent variable was attrition from recruit training. The attrition lists at the RTC, San Diego, were checked monthly and all attritees for whom we had reading data were noted, along with the reasons for attrition. Three predictor variables were used:

1. <u>Odds for Effectiveness (OFE)</u> — An estimate of the probability of completing the first tour in the service and being recommended for re-enlistment. This is an actuarial table used by recruiters in which a composite score is derived from years of schooling, number of expulsions and suspensions from school, and Armed Forces Qualification Test Score (Plag, 1968).

2. <u>Armed Forces Qualification Test (AFQT)</u> — This is a score derived from performance on the GCT, ARI, and MECH tests taken at the recruit station. Placement into a mental category is based on this score.

3. <u>Reading Ability</u> — Performance on the Gates-MacGinitie Reading Test described previously.

Results indicated that in our total sample of recruits, there was a 7.7% attrition rate as compared to an overall attrition rate at San Diego of approximately 10% during our test period. This disparity reflects the attrition of personnel prior to administration of the reading test, as well as personnel who were hospitalized at the time of testing and eventually attrited. The distribution of reading ability for the attritees and nonattritees in our total sample is shown in Figure 5. The primary reasons for attrition in our sample were inaptitude (academic and military), physical disability, and psychological disability. Approximately 30% of the attritees were in each of these three categories. Even with this wide variability in the reason for attrition, the data in Figure 5 indicate very different distributions of reading skills for attritees and nonattritees (median reading levels are 8.2 and 10.9 respectively). The probability of attrition in each of the reading categories is shown in Table 2. A clear and systematic relationship between reading ability and the probability of attrition is evident. The less-than-4th-grade readers have a 0.36 probability of making it through boot camp, while the probability is 0.96 that the above-10th graders will make it.

While reading ability is clearly related to attrition, the question still remains as to the independent contribution it makes relative to other available indices. To answer this question, we entered the reading, OFE, AFQT, and years of education scores into a stepwise multiple regression analysis with attrition, a dichotomous variable, as the dependent measure. The multiple regression analysis of the initial sample yielded reading as the strongest predictor of attrition ($r = 0.25$). The addition of the remaining three predictor variables added only 0.02 to the multiple correlation. The reading-attrition relationship increased to 0.33 in a cross-validation sample, and here the remaining three variables added only 0.003 to r. The independent correlations of OFE, AFQT, and years of education with attrition in the initial sample were 0.19, 0.17, and 0.14 respectively. The comparable correlations in the cross-validation sample were 0.16, 0.18, and 0.12.

In summary, these attrition data indicate that attrition rate systematically decreases as reading ability increases, and this holds through the full range of reading levels. Reading is the best single predictor of attrition from boot camp. The degree of relationship between reading and attrition is significantly larger than the correlation of attrition with those measures currently used by recruiters to predict the successful sailor. While this latter finding indicates that *reading contributes significantly to the prediction of attrition independently of the other variables*, we performed an additional analysis to confirm this conclusion.

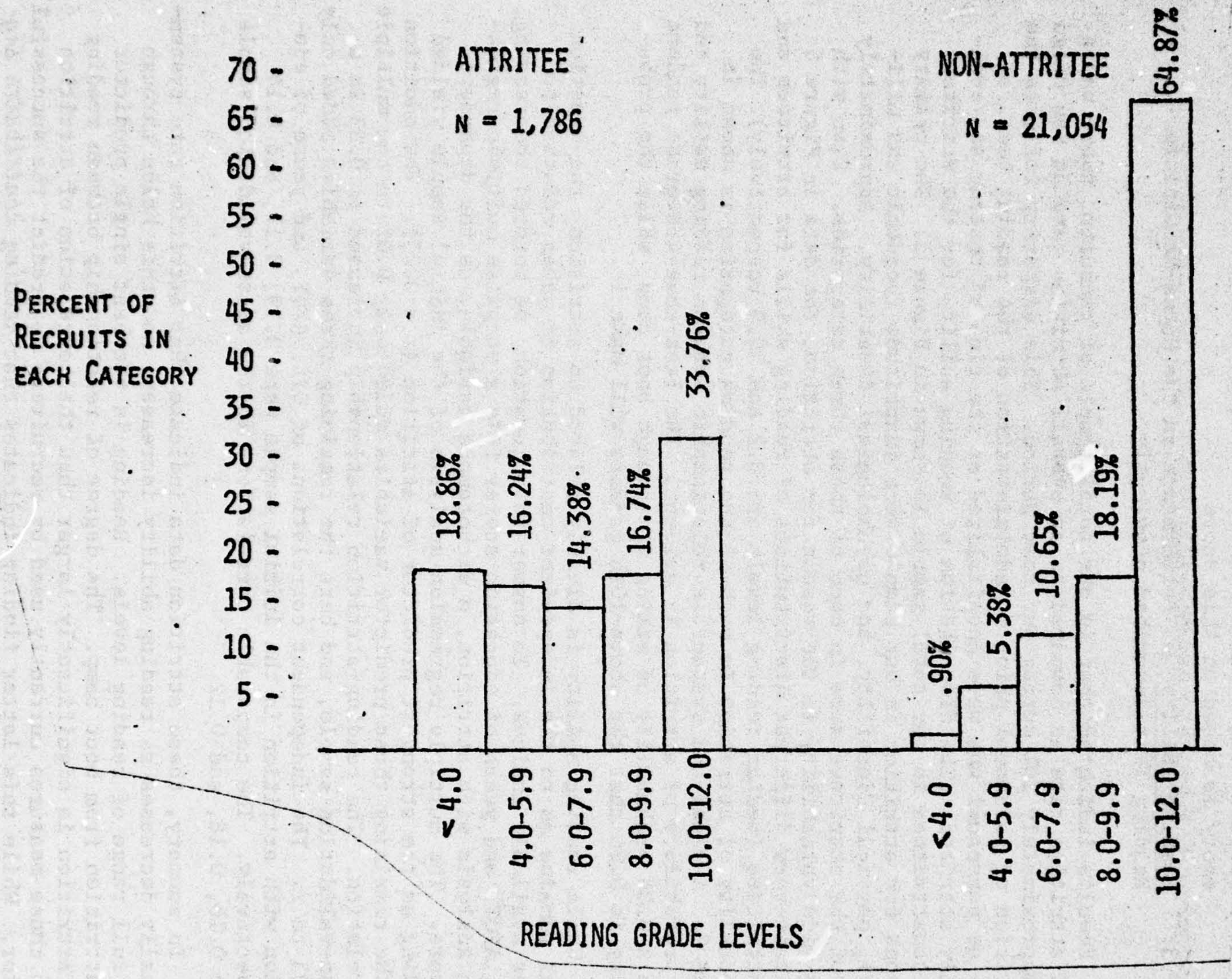


Figure 5. Reading Ability for Recruits Attriting & Graduating From RTC, San Diego, June 1974 to January 1975.

TABLE 2. SAN DIEGO ATTRITION RATE WITHIN READING CATEGORIES: JUNE 1974 — JANUARY 1975.

	Reading Grade Levels					
	< 4	4.0-5.9	6.0-7.9	8.0-9.9	10.0-12.0	Total
Number of men	526	1,424	2,500	4,129	14,261	22,840
Number attriting	337	290	257	299	603	1,786
Percent attriting	64.1%	20.4%	10.3%	7.2%	4.2%	7.8 %

A stepwise regression analysis was performed on the total data sample. This time, however, OFE, AFQT, and years of education were entered into the analysis before reading. These three variables together yielded a multiple correlation with attrition of 0.20. The addition of reading as a predictor raised this correlation to 0.30. That reading ability as measured by the Gates-MacGinitie test contributes to our prediction of attrition over and above other tests involving reading and non-reading components (AFQT), as well as amount of education, strongly suggests that it is reading skill with its various components (e.g., language, semantic knowledge) per se and not other abilities underlying reading test performance (such as perseverance to complete high school, conformity to expectations) which is related to attrition. Since the primary emphasis in boot camp is acculturation and does not involve excessive use of manuals, the relevance of reading skill to performance after boot camp should be even greater. With regard to selection policies in the Navy, these data indicate that use of a reading test would yield more accurate predictions of attrition than the measures currently in use.

Reading in Relation to A School Performance

The A School sample and the predictor variables are as described in the previous section. The dependent measure for this portion of the project was average performance on a weekly paper-and-pencil test. Subsequent analyses will examine an "in school" job performance measure, as well as the final score in the school, which is a score based on all previous testing.

Only a preliminary analysis of the relationship of each of our predictor variables to weekly test performance has been completed. In this analysis, we were again interested not only in the degree of relationship between reading and performance, but also how that relationship compared to the largest correlation between one of the current predictor tests and performance. The data relevant to these considerations is presented in Table 3, where it can be seen that the reading/performance correlations range from 0.20 to 0.50 across the six schools, while the maximum predictor test correlation with performance ranges from 0.21 to 0.61.

Additional analyses are planned using multiple regression techniques to further evaluate the usefulness of a measure of reading ability for predicting A School weekly paper-and-pencil test scores. If lower ability readers are poor risks in the A Schools, as suggested by the data of Table 3, an A School preparatory reading training program might be instituted to provide the deficient readers with the necessary skills for successful performance in their technical training.

TABLE 3. CORRELATION OF READING, PATTERN MATCHING, AND THE MOST PREDICTIVE TEST IN THE BASIC TEST BATTERY (BTB) WITH WEEKLY TEST AVERAGE IN SIX NAVY "A" SCHOOLS.

SCHOOL	WEEKLY TEST AVERAGE CORRELATION WITH: READING	PATTERN MATCHING	BTB TEST
HULL MAINTENANCE TECHNICIAN*	.20	.05	.21 (ARI)
SHIP'S SERVICEMAN	.26	.16	.21 (ARI)
MESS SPECIALIST	.43	.22	.32 (ARI)
QUARTERMASTER	.50	.48	.61 (ETST)
INTERIOR COMMUNICATIONS	.39	.39	.50 (ARI)
ELECTRICIAN'S MATE	.23	.16	.44 (ETST)

*A WEEKLY TEST IS NOT ADMINISTERED IN THIS SCHOOL AND THEREFORE THE DAILY TEST AVERAGE SCORE WAS USED

LITERACY TRAINING

I would like now to narrow the focus of consideration to personnel demonstrating major deficits in reading skills — the below-6th-grade reader. The Marine Corps and the Navy provide reading training during recruit training for these low literate men. The objectives of the training are (1) to provide a level of literacy skill to all personnel so as to ensure fleet effectiveness and fleet safety, and (2) to provide the literacy skills necessary for equal opportunity in attaining upward mobility and a successful career (Steward, 1974; Academic Remedial Training, 1975). Clearly, data on the reading requirements in the Navy presented earlier and described in more detail by Carver (1973b) and Curran (this conference) suggest that the training criterion, roughly a 5th-grade reading level, is not adequate to achieve either of these objectives. Unfortunately, the data required to evaluate whether or not the objectives are being met has not been gathered. However, there is data relevant to the more immediate objective of the literacy training; that of improving reading skills. In this section, I will describe the effectiveness of three literacy programs in achieving this more immediate objective. In addition, I will describe three experimental training programs which have just been initiated. Finally, I will describe my own research efforts to characterize the reading deficiencies and to predict the probability of success in the current training programs.

Current Programs

The three programs to be described are Navy programs in San Diego, California, and Orlando, Florida, and a Marine Corps program in San Diego, California. Program descriptions and evaluation data were obtained from the following sources: Marine Corps (Stewart, 1974, 1975); Navy in San Diego (Academic Remedial Training, 1975); Navy in Orlando (Research for Better Schools, 1974).

The Orlando program was an experimental effort to test the Individualized Learning for Adults package developed by Research for Better Schools (RBS Program). This is a self-paced program for illiterate adults and provides training material to the 9th-grade level. Training was on general literacy and focused on the specific skills of recognition of sound-symbol relations (phonics), word attack (vocabulary), comprehension (literal and interpretive), and study methods. There are 129 performance objectives in the program with 32 of the objectives devoted to decoding skills, 21 to word attack, and 41 to comprehension skills. Five entrance tests were used to determine the appropriate starting level for an individual and performance during training was regularly assessed using 70 different pre- and post-tests. The training materials were 129 study booklets and 54 audio cassettes specially prepared for this program. The maximum student-instructor ratio was 7 to 1 (4 instructors and a maximum of 28 students in a classroom). The instructors were Navy men selected on the basis of having a college degree (area of study unspecified). Instructional time ranged from 10 to 160 hours with a mean of 60 hours.

A total of 43 male recruits participated in the RBS program between September 1973 and February 1974. The criterion for selection was a score of 7.0 RGL on the combined vocabulary and comprehension subtests of the Gates-MacGinitie Reading Survey D. The methods of selecting recruits to test were not specified. After three weeks in the program, the student was retested in an alternate form of the Gates-MacGinitie test. If this score exceeded a 7.0 RGL, the man was returned to recruit training; if under 7.0, he continued through the entire RBS program.[2]

The Marine Corps (MC) program in San Diego was devised especially for the Marines in conjunction with the local adult school. This is the only program involving civilian instructors, all of which are female and hold teaching credentials and adult school certificates. The program involves some self-pacing, but is classroom-based and fixed at four weeks (120 hours) of instruction. The training is in general literacy and instructional time, according to the program syllabus, is distributed as follows: decoding (phonics), 40%; reading speed and comprehension, 42%; vocabulary development (sight vocabulary and word attack), 15%. Thus, there is considerable emphasis on decoding skills in this program. A wide variety of commercially available training materials and diagnostic tests are used in the program, including the Controlled Reading Skill Development Series (Educational Development Lab), McCall-Crabbs series (Columbia University), Reading Attainment System (Grolier Education Corp.), Reading for Understanding (Science Research Associates), and Phonic Word Blend Flip Charts (Kenworthy Educational Service).

All recruits at the Marine Corps Recruit Depot in San Diego who read below the 4.5 grade level are accepted into the MC program. Initial screening is accomplished through the administration of the Gates-MacGinitie Reading Survey D to all recruits scoring below approximately the 70th percentile on the general classification test. Recruits scoring below the 4.5 RGL are retested on an alternate form of the Gates-MacGinitie and only if they once again score less than a 4.5 RGL are they admitted to the program. The score on the second test serves as the person's entrance score. Evaluation data are available on 492 recruits entering the program between February 1974 and July 1975.

The Navy program in San Diego (NSD program) is the only program of the three under discussion which is run entirely by the service. Instructors are Navy personnel having a college degree and the training materials are selected or developed by these instructors. Approximately 25% to 30% is devoted to phonics training. The phonics materials are based on the Motts Phonic System (unreferenced) and are similar to those used in the MC program except that phonics is concentrated in the first week of training. Successive weeks of the program deal with vocabulary (60 words per week), comprehension, and reading speed. The program is lock step and is a minimum of 3 weeks. Failure in any week results in repeating that week. Thus, the program averages 4 weeks (92 hours) and ranges from 3 to 6 weeks.

[2] The fact that a student was not retested until 3 weeks into the program appears inconsistent with the report that some men completed the program in 10 hours. No explanation is available.

Beginning in July of 1974 (FY 1975) specific testing for knowledge of sound-symbol relations was instituted. If skills in this area were adequate, the recruit bypassed the first week. Thus, in FY 1975 the average program duration was reduced to somewhat less than three weeks (69 hours) with a range of two to five weeks.

The training materials in the NSD program are a mixture of instructor-generated worksheets and commercially available supplemental materials. Hardware is limited to reading pacing devices. Prior to FY 1975 the supplemental materials consisted of the McCall-Crabbs readers and a few magazines (e.g., Scholastic Magazine). In FY 1975 a wide variety of commercially available training packages, books, and magazines were introduced into the program. It is estimated that in both FY 1974 and 1975 approximately 70% of the vocabulary training material and more than 90% of the comprehension training material focused on general literacy.[3]

The NSD program is directed at recruits with an RGL between 3.0 and 5.5. Prior to FY 1975 all recruits failing the first academic test in recruit training were administered the Gates-MacGinitie Reading Survey D. If a recruit scored below the 3.0 level he was recommended for discharge. Recruits between a 3.0 and 5.5 RGL were admitted to the NSD program. In FY 1975 the Gates-MacGinitie test was administered to all recruits. Those scoring less than 5.5 RGL were then retested on an alternate form. Only those scoring between 3.0 and 5.5 RGL on the retest were admitted into the program.

The distinction between FY 1975 and FY 1974 procedures is important in the consideration of the effectiveness of the programs. The RBS program and the NSD program prior to FY 1975 admitted students on the basis of a single test administration. Since only low scoring individuals were admitted into the program and later retested, the change scores should be subject to considerable upward regression. The MC program and the FY 1975 NSD program, by selecting students after two test failures, should reduce this regression effect. Thus, the regression effect should increase the amount of artifact in the gain in the FY 1974 NSD and RBS programs to a greater degree than in the other two programs.

In addition to the program characteristics described thus far, differences in student characteristics and program policy may be expected to affect gains in reading within the programs. Tables 4 and 5 summarize the racial and educational characteristics of recruits entering each of the programs. The racial distribution is about equivalent in the MC and NSD programs except for an increase in the proportion of Malaysians (and thus, second language training requirements) in the FY 1975 NSD program.

---

[3] The NSD program is currently undergoing another revision. The new materials are estimated to be 70% to 80% Navy-related and the instruction will be individually paced.

TABLE 4

PERCENT OF RECRUITS IN EACH RACIAL CATEGORY FOR FOUR READING TRAINING PROGRAMS

	RACE			
PROGRAM	CAUCASIAN	BLACK	MALAYSIANS	OTHERS
RBS *				
MC	45	29	11	15
FY74 NSD	44	26	12	18
FY75 NSD	40	20	26	14

* RACE DATA UNAVAILABLE FOR THIS SCHOOL

TABLE 5

PERCENT OF RECRUITS AT EACH LEVEL OF EDUCATION FOR FOUR READING TRAINING PROGRAMS

	EDUCATION LEVEL				
PROGRAM	8	9	10	11	12+
RBS	7	16	26	30	21
MC	11	21	25	25	18
FY74 NSD	2	12	16	25	45
FY75 NSD	3	6	15	23	53

Educational differences, however, are considerable. Approximately 50% of the NSD students have a high school diploma, while only 20% of the students in the other programs are graduates. If a higher level of educational achievement is indicative of the acquisition of better developed study skills, then the NSD programs may show greater reading gains due simply to the relatively greater level of eductional achievement of their students. Parenthetically, it is interesting to note that the data in Table 5 reflect the conclusion drawn earlier that a high school diploma does not ensure that a person is functionally literate.

Table 6, while summarizing instructional gains in the programs, also contains information relevant to our consideration of the relative gains one might predict for the programs. First, the entry level reading scores at MC are considerably below those in the other programs. The values were not subjected to statistical comparison because of lack of adequate raw data and because of the widely varying Ns. However, the entry scores fall within that range of ability at which phonic skills are acquired. Thus, a 1 RGL difference in entry levels should reflect a meaningful difference in phonic skills, which are considered by instructors as the most difficult and time-consuming skills to acquire. In a post hoc vein, our own data and data from the MC and NSD programs indicate that entry reading level is a good predictor of amount of gain (a positive relationship). Therefore, on the basis of entry level scores, the MC program should yield smaller gains. Table 6 also indicates that the NSD programs attrite a considerably greater proportion of students. Since only graduates are entered into the post-test calculation, this greater attrition is an artifact which would increase the gain in reading skills found in the NSD programs.

Table 7 summarizes those program characteristics which might be expected to benefit the amount of gain found in each program. The RBS program may be characterized as "program" oriented, since it involved carefully developed training materials and procedures. The MC program, in contrast, capitalizes on instructor capabilities and extended contact with the instructors (a relatively long training period). Finally, the NSD program benefits are student based in that better students (higher RGL and level of education) enter the program and there is a more liberal policy of attriting students not showing progress or students demonstrating improper attitudes.

The gain scores and post-test scores presented in Table 6 indicate that despite considerable differences in program orientation, all of the programs yield approximately the same amount of gain and terminal level of reading performance. Potentially, training of some of the men in the RBS program was terminated after 3 weeks due to their achieving a 7.1 RGL or higher on a test. However, all other men as well as the men in the other 3 programs were required to complete a fixed course of instruction. Additionally, in all programs, except perhaps the FY 1974 NSD program, training materials went well beyond the sixth-grade level. Thus, the similarity between programs in the final level of achievement does not

TABLE 6. MEAN READING GRADE LEVEL* ON PRE- & POST-TESTS, GAIN IN READING, AND ATTRITION RATE FOR FOUR READING PROGRAMS.

	READING SCORES			ATTRITION
	PRE TEST	POST TEST	GAIN	
RBS (N=43)**	4.5	6.3	1.8	.00
MC (N=490)	3.5	5.7	2.2	.10
FY74 NSD(N=785)	4.2	6.0	1.8	.22
FY75 NSD(N=658)	4.2	6.1	1.9	.24

* READING WAS ASSESSED ON THE VOCABULARY AND COMPREHENSION SUBTESTS OF THE GATES-MACGINITIE READING SURVEY D.

** THE "N" DOES NOT REFLECT ATTRITION FROM THE PROGRAMS

TABLE 7. PROGRAM CHARACTERISTICS WHICH ARE EXPECTED TO RESULT IN RELATIVELY GREATER REPORTED GAINS IN READING.

	PROGRAM			
CHARACTERISTICS	RBS	MC	FY74 NSD	FY75 NSD
GREATER INSTRUCTOR QUALIFICATIONS		X		
LONGER PROGRAM DURATION		X	X	
GREATER TRAINING MATERIALS EXPENDITURE	X			
GREATER INDIVIDUALIZATION	X			
HIGHER ENTRY READING LEVELS	X		X	X
HIGHER EDUCATIONAL LEVELS			X	X
HIGHER RATE OF ATTRITION FROM PRE TO POST TEST			X	X
GREATER STATISTICAL REGRESSION FROM PRE TO POST TEST	X		X	

seem to be due to a ceiling effect created by a commonality in training criteria. The difference between the programs in terms of amount of gain is only a .4 RGL. Considering the substantial differences in characteristics between the programs, this small of a difference in effectiveness suggests the conclusion that *within the confines of a short duration, instructor-based, low literacy program, the amount of instructional gain in general literacy is fixed and determined by the simple exposure to the educational setting.* This conclusion is further supported by the 1.2 RGL gain in general literacy obtained in the Army's Functional Literacy Training program (Sticht, Caylor, Fox, Hauke, James, Snyder, & Kern, 1973). This program, which underwent considerable experimental development, employs qualified teachers, is highly individualized, and focuses on functional reading skills. Nonetheless, it possesses the defining characteristics for my conclusion and the gain in general literacy is comparable to the gains obtained in the naval service.

Results obtained by Shennum, Aiken, and Thomas (1975) suggest that this invariance in reading gains may even apply to specific instructional procedures. These investigators examined three procedures for increasing reading speed for recruits in the final week of the FY 1975 NSD program with the hypothesis that excessively slow reading speeds (frequently found in these programs) may be an important contributor to poor comprehension. The training method of primary interest involved simultaneous reading and listening with the rate controlled through a variable rate speech compressor. Comparison conditions were (1) simply listening to the compressed speech, and (2) reading with no listening but with rate goals specified and feedback given. Recruits in these conditions spent two hours per day for six days reading a 40,000-word novel written at the 8th-grade level. A pre-test/post-test evaluation using different materials indicated that each condition produced an increase in reading rate of about 60 words per minute and a 20% increase in comprehension. These gains, while constant across conditions, were nonetheless due to training, since a no-treatment control showed no gain in rate or comprehension over the six-day period.

The general conclusion I have drawn from these evaluations should not lead to the assumption that literacy training programs cannot be improved. A consideration of alternative criteria for training success (the criteria are currently amount of gain and exit level in general literacy), a shift to training of men with less deficiency, or the introduction of refresher or retraining programs may all yield more effective reading programs. Perhaps the most important consideration is a rethinking of the criteria for training success. In most Navy training, program effectiveness is heavily weighted by the time required to reach criterion. The time factor is very important, since students and staff are all on salary. If a time criterion were applied to the reading programs, it is highly probable that individualized, computer-based programs employing mastery learning techniques would be by far the most cost-effective (Atkinson, 1972; Ball & Jamison, 1973; Bloom, 1974). The only time information available on the present programs comes from RBS, where it was found that on the average,

41 hours of instruction was required for a one-year gain in reading. However, that one-year gain required 52 hours of instruction for recruits without decoding skills, but only 25 hours when recruits could decode.

The criterion for literacy training should also be rethought in terms of the ultimate goals of the program. The goal is not to produce more literate men, but rather, to give them the skills necessary to effectively perform the reading tasks required of them in the service. These skills involve an ability to find information readily in manuals, to read and comprehend procedural directions and instructions, to determine the relative importance of information on a system, etc. The vocabulary requirements and the reading strategy required in these tasks differ considerably from the requirements in general literacy. Tests based on service-relevant reading material and tasks will very likely show program gains not reflected in general literacy tests.

Still related to the ultimate criterion of increased Navy effectiveness, a program addressing men with minor reading deficiencies may be expected to yield greater gains in performance effectiveness as compared to low-literate programs even if both programs produced the same gain in the same time. While the advantages of the former program were discussed earlier, no information was available on the relationship of reading improvement to entry reading level. Data derived from these literacy programs indicate that in addition to the gains in Navy effectiveness, a program for moderately literate men would require less training time to achieve the same gains in reading. The NSD personnel refuse to accept less than 3.0 RGL readers as they have found training of these men to be virtually impossible. RBS reports that when a man is not proficient in phonic skills, he requires twice as much training to produce a year gain in reading. In my own work at NSD, I have found a strong linear relationship between entry level and gain ($r = .35$ and $.59$ in two samples of native English speakers).

Literacy training programs have frequently been criticized for the tendency to be one-shot programs. Clearly, meaningful permanent gains in reading cannot be expected in a three-week program. Exercising of the new skills and refresher training is necessary. The NSD personnel retested 58 graduates from the FY 1975 program two to seven weeks after graduation. The delay between the final test and retest did not produce any systematic differences in change scores, which indicates that the loss due to termination of training occurs within two weeks. The average graduation score for these 58 men was 6.3 RGL, while the average retest score was 5.1; a 1.2 RGL loss in reading skill due to the termination of training. If these men were representative of the FY 1975 input into the program, these data indicate a .9 RGL permanent gain from the entry reading level of 4.2 RGL. However, this gain is still subject to some regression effect, and so the permanent gain in the program is likely a .5 to .8 RGL. While retest data is not available on the other Navy programs, there is no reason to suspect the permanent gain would be any greater than that obtained at NSD.

The exit reading score indicates that the men have the capability of reading above the level obtained in retesting. To achieve that higher level on a permanent basis, however, will require mini-refresher courses over an extended time frame.

The fact that the reading programs produce approximately equivalent gains and that there is considerable loss in skill after a man leaves the program should not detract from their relative effectiveness. The pre-test and post-test gain of approximately two years in these compulsory and time-compressed programs is indeed significant. Looking at the gains in another way, the men, on the average, ranked at the 30 percentile for 5th-grade students at entry into training, while on the post-test, the men were at the 69 percentile. Thus, during the course of approximately four weeks of training, these men move from the lower to the upper third of the distribution for this particular reference group. The 2 RGL gain is also significant relative to an average .5 RGL gain achieved in civilian adult school programs of the same total instructional hours, but spread over five months (and therefore, providing distributed practice) and offered on a voluntary basis (Kent, 1973). Similarly, experimental reading programs for school children at the same reading skill level as the recruits have produced less than a year gain in a year of instruction (Battelle Institute, 1972). In comparison, the naval service programs have proved exceedingly effective in improving reading skills. My discussion, instead, addresses the questions of whether more effective programs can be developed and whether the programs produce gains to the service beyond the increase in reading.

With regard to the effects of the reading programs on later Navy effectiveness, graduates of the NSD and MC programs have been tracked through boot camp to determine their rate of attrition. The success of a sample of 301 MC program graduates was assessed 3 to 14 months after graduation. The attrition rate of this sample was 12% as compared to an overall boot camp attrition rate of 10%. Since boot camp is only 11 weeks, these data indicate that graduates of the MC program have an attrition rate equivalent to or less than the Marine Corps average. Additionally, only 33% of the graduates were judged to have a promotion rate slower than normal, while for 17% the rate was judged to be faster than normal. Unfortunately, an untrained sample is not available for comparison. (Perhaps most men are typically *judged* to be progressing faster than normal.) The data do suggest that the MC program graduates are progressing through the service at a normal rate. However, we do not know to what degree reading ability is predictive of performance in the Marine Corps. (It is evident that the reading requirements are less than in the Navy.) Thus, the "normal" performance of these graduates may be due to some characteristic of the reading program or it may be due to the lack of any relationship between reading ability and service effectiveness.

Boot camp attrition rate for a sample of 387 graduates from the FY 1974 NSD program was 12%. This can be compared to an average 10% attrition rate at the boot camp and an attrition rate of 18.7% for fourth- and sixth-grade readers. As with the MC program, the graduates of the NSD program appear to perform as well as the average recruit. Here, however, we have data pointing to a relatively strong relationship between reading and attrition (see Table 2) and thus, the 12% attrition rate for graduates appears to be of meaningful significance.

The graduates of both the MC and NSD reading programs appear to be performing as average recruits. However, a basic question is whether their level of performance in the service is due to the improved skill level and attitude obtained from the reading program. While a definitive answer to this question is not possible, it seems highly unlikely that 1 RGL increase in ability would substantially affect the performance level of a functionally illiterate recruit. A more likely effect of the program is that they serve to filter out recruits having low capability for learning or having a poor attitude. These people are represented in the 12% and 14% attrition rates in the MC and FY 1974 NSD programs respectively. The only way we are going to be able to adequately assess the effectiveness of reading training on later service effectiveness is through an adequate experimental design in which only a portion (randomly selected) of the personnel eligible for reading training are actually assigned to the program and the program resulted in meaningful gains in reading skill. A judgement of reading training effectiveness would then be based on attrition rates from the point of assignment, rather than examining only the graduates of the reading program.

Experimental Reading Programs

In addition to the ongoing programs for recruits, there are three experimental programs under Navy sponsorship. Since these programs are only in the initial stages of development, I will only briefly describe them. Hornbeck and Montague, under the sponsorship of the Center, are developing programs for training phonic, or decoding, and vocabulary skills in the low literate recruit. The training is computer based and has as dual goals, the evaluation of training procedures and the evaluation of new computer hardware and software systems. The interactive computer system has graphic capabilities allowing the student to "point" to stimuli, e.g., to syllabicate a word shown on the screen. More importantly for the low literate training, the terminals are equipped with a Vortax voice synthesizer. This computer-generated speech capability permits a fully interactive and individualized training program with computer presentation in either or both the auditory and visual mode. A programming language incorporating the graphic and voice synthesizer capabilities has recently been generated and the courseware for the phonic skill training is nearing completion.

The second experimental program is being developed by Carver (1973d) under the sponsorship of the Office of Naval Research. The training in this program, rather than focusing on specific reading skills, will involve 150 hours of practice on the terminal objective — reading and comprehending prose material. Prose passages will be presented over a PLATO computer terminal, thus permitting individualized instruction. The passages will be presented in a modified cloze (Taylor, 1957) format in which every fifth word is deleted and the student makes a multiple choice selection to fill the blank. This procedure forces continued attention to the passage and provides constant monitoring of performance. The primary objective in Carver's work is to determine if there are two types of poor readers — those lacking the intellectual capacity, and those lacking the educational experience. The latter are expected to gain from reading training. Thus, students will be screened for intellectual ability using the Raven's Progressive Matrices and high and low Raven's scorers will be compared. Those receiving training will consist of 4th-grade level readers. Pilot testing of this system has recently begun.

The final experimental program involves the evaluation of the Encyclopedia Britannica program for functionally illiterate adults. The program is at a considerably lower level than the other programs we have discussed and assumes no reading ability. Training is in pronunciation, listening, and basic reading skills. This program has been instituted aboard three ships and participation is voluntary. The experimental evaluation of this program, by the Chief of Naval Education and Training Support, is just getting underway.

Prediction of Training Success

As a final topic in the review of the Navy's literacy research, I would like to present some of the results obtained in my own research efforts to characterize the successful student in the NSD reading program. For the past several months, we have been administering a battery of tests to all recruits entering the program and examining test scores in relation to the amount of gain[4] in reading ability. Our first sample consisted of 111 recruits, 32 of whom had English as a second language. We then made some modifications in the test battery and have recently begun collecting data on a second sample. This sample at present consists of 41 recruits, 21 of whom have English as a second language. It should be borne in mind that the data I am presenting applies to the 1975 NSD program which was described previously. Programs differing in instructional strategies would likely produce different relevant variables.

One of the first hypotheses we examined was that the remediability of poor readers is dependent upon their intellectual ability. This is a basic hypothesis in Carver's (1973d) experimental reading program discussed previously. To test this hypothesis, we used the pattern matching (PM) test (derived from the Raven's Standard Progressive Matrices as described previously) as a part of our battery. The split-half reliability (corrected for attenuation) of the PM, based on the scores of 1,200 recruits, is .85. Our dependent variable was the gain in reading scores over the course of the reading program. If a student was attrited from the program, he was given a score of 2.6 — .4 RGLs below the lowest exit score. Mean PM performance was consistent across the four groups (successive samples crossed with English as a first or second language) ranging from 19.5 to 22.5 (σ = 4.9 to 6.6). This is significantly below the mean of 28.3 obtained for all recruits at the San Diego boot camp. The mean gain in reading ranged from .8 to 1.8 RGLs, but was not consistently related to a language group. The gains and the correlation of gain and PM performance is presented in Table 8 for the English as a native language (E 1st) and English as a second language (E 2nd) groups.

The correlation between PM and gain is low and not statistically significant ($p > .05$) across all four groups. If anything, there is a tendency toward a negative correlation between gain and pattern matching for the E 1st group. Thus, intellectual ability of low literate recruits is not predictive of their gain in this reading program. This result is consistent with our finding with a random sample of all recruits, where it was found that PM was not strongly related to general reading ability (r - .43). It seems that while intellectual ability is a component in reading skill, it alone does not set the limits for the level of skill attained in a reading program, at least as found in our heterogeneous adult population.

---

[4] Analysis involving post-test scores have yielded comparable results.

TABLE 8. MEAN (STANDARD DEVIATION) READING GAINS* & THE CORRELATION OF GAIN WITH PREDICTOR VARIABLES FOR NATIVE LANGUAGE (E 1st) & ENGLISH AS A SECOND LANGUAGE (E 2nd) GROUPS IN TWO SAMPLES.

		CORRELATION OF GAIN WITH:		
	GROUP (N)	PATTERN MATCHING	DECODING SKILL	MEAN GAIN
SAMPLE 1	E 1ST (79)	-.02	.40	.8 (1.7)
	E 2ND (32)	.19	.30	1.8 (1.9)
SAMPLE 2	E 1ST (21)	-.39	.54	1.23 (1.49)
	E 2ND (29)	.18	.28	1.0 (1.1)

*READING GAIN IS CALCULATED AS THE POST-TEST/PRE-TEST DIFFERENCE ON ALTERNATIVE FORMS OF THE VOCABULARY AND COMPREHENSION SUBTESTS OF THE GATES MacGINITIE READING SURVEY D.

Our second hypothesis was that those students more proficient in decoding printed words into phonological units would show greater gains in reading. This is an often reported finding by the instructors in the reading program and in the NSD program it is the rationale for rejecting recruits below the 3.0 RGL. The word knowledge subtest of the Wide Range Achievement Test was included in our battery to test this hypothesis. The word knowledge test involves the presentation of a graded series of printed words with the subject's task being simply to pronounce the words correctly. Correctness of pronunciation is a test administrator judgement and grading is liberal with respect to regional and ethnic differences in pronunciation. The test does not take into account the student's knowledge of the meaning of the words.

Results obtained with the decoding skill test supported the hypothesis. Performance on this test was consistently one of the best predictors of gain, with the correlation being somewhat higher for the E 1st sample ($r = .40$ and $.54$) then for the E 2nd sample ($r = .28$ and $.30$). Thus, as indicated by the instructors, decoding skill is an important determinant of ability to learn to read in a short duration program. Decoding skill, however, is clearly not the only factor. The mean decoding performance (in grade levels) for the E 2nd samples was 10.1 as compared to a mean of 4.7 for the E 1st, a difference of 2 standard deviations. The magnitude of this difference was consistent across samples, yet the gain in reading for the E 2nd group is not consistently greater than for the E 1st group. We feel that a second major distinction between the language groups, which will account for differences in gains, is semantic knowledge. The E 2nd consists primarily of Filipinos who have been taught decoding rules for English since grade 1. However, their experience with the language has been largely limited to schoolroom use. Thus, while they can readily decode the printed word, their semantic knowledge is deficient relative to the E 1st and will affect their progress in the remedial program.

We included reading and listening tests in our battery to test the above hypothesis, as well as to test the hypothesis that a large difference between reading and listening skills is predictive of the gains to be made in a remedial program. The tests were developed from the vocabulary subtest of the Nelson (1962) reading test. The odd-numbered items on the test formed our reading test and the even-numbered test items formed the listening test. Presentation of reading and listening items was alternated in blocks of five. "Listening" test is somewhat of a misnomer, since both reading and listening items were presented via slides at a 12-second rate. For the listening items, however, the stimulus and alternative were read to the subject while he viewed them. Split-half reliability of the listening test calculated on our first sample yielded an $r = .79$ (corrected for attenuation).

The mean listening and reading performance for each language group is presented in Table 9. The data indicate, as might be expected, that E 1st recruits have a larger listening than reading vocabulary. Their listening skill is exercised constantly through interaction in an English language community, while their reading skill training is likely to be limited to use in the educational system. The E 2nd, on the other hand, have equivalent reading and listening vocabularies indicative of an exposure to English, spoken or written, which is limited to the classroom. The data further indicated that the E 2nd group has a vocabulary knowledge greater than the reading vocabulary of the E 1st group, but less than the E 1st listening vocabulary. The E 2nd, while being able to decode far more words than the E 1st group, have more limited semantic knowledge of those words.

We have calculated the correlations between reading and listening skills and gain scores. However, the pattern of correlations is not stable or clearly interpretable. Thus, I will delay discussing that data until our second sample is complete. We have included a number of additional measures in our battery for the second sample. These include some of the traditional measures of Navy effectiveness, such as number of dependents and years of education, as well as tests of immediate memory and ability to maintain a high level of performance at a tedious task. The intent of these tasks is to look for what may be general ability deficits in our low reading ability recruits.

TABLE 9. MEAN (STANDARD DEVIATION) READING AND LISTENING ABILITY SCORES FOR NATIVE LANGUAGE (E 1st) AND ENGLISH AS A SECOND LANGUAGE (E 2nd) GROUPS COLLAPSED ACROSS TWO SAMPLES.

	READING	LISTENING
E 1ST	26.1 (5.1)	31.1 (5.0)
E 2ND	29.7 (5.0)	28.0 (4.8)

COMMENTS ON THE PAPER BY DUFFY

Burl Gray
Behavioral Sciences Institute

The comments by Gray focused on the parallels between the concern shown in Duffy's paper for understanding the characteristics of the reader, and those same kinds of concerns so frequently found in the civilian education world. Gray pointed out that he thought too much attention was given to understanding the learner, and not enough attention was given to describing the instructional method that students are exposed to. He suggested that the only way one will be able to accurately estimate reading gain as a function of time is to have a teaching method highly specific to some goals or objectives; then one can make a meaningful statement about gains per hour of instruction. In addition to making predictions of success in a program more accurate and meaningful, having a clearly defined method and procedure is useful when investigating failures from that program. In this case, one can make definite adjustments to the program until the failure rate is minimal. Gray closed his remarks by reiterating that in the military work, more attention needed to be given to specifying instructional methods and procedures. Continued worry about the characteristics of the reader may imply that the fact that he does not read is *his* problem, so trainers may not be too motivated to improve *their* work, and the problem may be expeditiously solved by discharging the person or changing his job assignment, or some other change in category rather than competency.

ONGOING R&D IN ARMY LITERACY TRAINING

John S. Caylor
Human Resources Research Organization

The Army's problem in dealing with recruits of low level reading and language skills is a recurrent one. In the past, the response to this problem has typically been the initiation of brief, limited, stop-gap programs of general literacy training whose content and objectives bore little demonstrated relationship to the requirements of Army job training and job performance. The most recent recurrence of this problem arose in 1966 with Project 100,000, which brought large numbers of previously ineligible low aptitude men into the Army. At this time, the Army undertook a program of research and development which was completed in 1975. This paper describes that program of literacy research and development activities, and briefly mentions the now ongoing research.

As an immediate response to the influx of marginally literate personnel under Project 100,000, the Army established Army Preparatory Training (APT). The goal of this program was to bring recruits reading below the fifth-grade level on a standardized reading test to that criterion before their entry into basic training. With widely varying local options, this program provided general educational development training in reading and arithmetic for a period of not more than six weeks, at which time all students were advanced to Basic Combat Training regardless of their reading achievement.

Shortly after the APT program was established, the Army, under contract with HumRRO, undertook a systematic long-range program of literacy research and development projects which had two major objectives:

1. To determine and define the nature of literacy problems in the Army by studying the actual operational literacy demands of Army jobs in conjunction with the literacy skills of Army personnel.

2. To develop a literacy training program to provide job-demanded functional literacy skills to meet the reading requirements of Army jobs.

Detailed descriptions of the projects undertaken to meet these goals have been reported elsewhere (Sticht, 1975c). The earlier phases of this program were primarily of a research nature and these will be summarized in terms of the work to determine the functional reading requirements of Army jobs and the reading ability level of the Army personnel available to do these jobs. The latter phases of this program were more heavily developmental; those phases will be described in terms of the job functional literacy training program which was produced, its operational

effectiveness under Army-wide implementation as a front-loaded program conducted before job training, and the feasibility of integrating this job reading training program with job skills training.

DETERMINATION OF THE LITERACY DEMANDS OF JOBS

To determine the literacy requirements of Army jobs, three different approaches were used. In these three approaches, the reading requirement of a job was established in terms of one of the following:

1. Direct measures of job knowledge and job performance.
2. The readability (reading difficulty level) of the Army manuals prescribed for use in learning and in doing the job.
3. The specific job reading tasks inherent in performing the job.

Each of these approaches is described below.

Job Proficiency

Our first approach to determining the literacy demands of jobs was to examine the relationship between the general reading ability level of job incumbents and their job proficiency. Job proficiency was measured both by a written test of job knowledge and by an extensive hands-on sample of job performance. For each of four Military Occupational Specialities (MOSs), the literacy requirement of the job was estimated as the lowest reading grade level at which no more than a chance proportion of men fell in the bottom quartile on the job proficiency measures. These analyses indicated a consistent relationship between literacy and job proficiency, and suggested the requirement of 7th-grade reading level for Cooks, 8th-grade reading level minimal requirement for Armor Crewmen and Vehicle Repairmen, and 9th-grade reading level for the Supply Clerk's job.

This approach to establishing the reading requirements of jobs has the advantage of using job proficiency measures directly as criteria. This work suggested that different jobs do have different levels of literacy demands and that reading requirements for these four high-density jobs lay well above the levels typically set as objectives for remedial reading training programs. There are drawbacks to this appraoch. Clearly it is prohibitively expensive to obtain hands-on job proficiency measures in a variety of jobs. A different problem arises from the job proficiency measures themselves, for they represent the resultant effect of many factors, of which literacy is only one. Since the job proficiency approach totally ignores the job reading materials themselves, we then turned to the study of those reading materials used in training for and performing the job.

Readability Approach

Our second approach to determining job reading requirements was to study the readability of reading difficulty level of Army manuals used on various jobs. The FORCAST readability index was constructed to estimate the reading grade level of ability needed by the adult Army population to read and comprehend technical job reading materials (Caylor, et al., 1973). FORCAST estimates of the readability of manuals indicated these results:

1. More than half of the job manuals in each of seven jobs exceeded the 11th-grade level of reading difficulty.
2. The average readability level of the materials in each of these jobs far exceeded the average reading ability of men working in these jobs.

Table 1 shows that although there are clear differences in readability of job printed materials among the MOSs, *all* the MOSs show readability levels well above the 9th-grade level.

TABLE 1

Cumulative Percentage Distribution of Job Reading Materials for Seven MOSs and Seven FORCAST Readability Levels

RGL[a]	Military Occupational Specialty[b]						
	11B20 (N = 104)	26D20 (N = 95)	63B20 (N = 108)	71H20 (N = 95)	76Y20 (N = 83)	91B20 (N = 90)	95B20 (N = 138)
6-6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7-7.9	1.0	0.0	0.0	0.0	0.0	0.0	0.7
8-8.9	4.8	0.0	3.3	1.1	3.6	2.2	5.0
9-9.9	18.3	4.2	13.4	2.2	10.8	24.4	15.1
10-10.9	41.4	9.5	36.3	3.3	20.4	47.8	34.0
11-11.9	71.2	42.1	61.8	37.0	57.6	77.8	62.2
12.0+	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[a]RGL—Reading Grade Level of difficulty of job printed materials determined by the FORCAST formula.

[b]11B20, Light Weapons Infantryman; 26D20, Ground Control Radar Repairman; 63B20, Wheel Vehicle Repairman; 71H20, Personnel Specialist; 76Y20, Armorer/Unit Supply Specialist; 91B20, Medical Specialist; 95B20, Military Policeman.

The readability technique offers a low cost method for estimating the overall reading demands of job manuals. However, it does not provide a direct indication of how well men can read and use their manuals for the *reading tasks performed on the job*; for this information, we need to test people on samples of job reading tasks using job reading materials.

Job Reading Task

Our third general approach to assessing job literacy requirements consisted of studying directly the relationship between general literacy skill and performance on job reading task tests; that is, reading tests constructed of actual job reading materials used in performing actual job reading tasks. Job reading tasks were identified by means of structured interviews with job performers at their work location. Job performers were asked to report instances of their use of printed materials in performing job tasks, to describe the information they sought to perform the job, to obtain the printed material, and to indicate the needed job information in the manual. These verified job reading tasks were then structured into Job Reading Task Tests which were standardized and normed on Army samples whose general reading ability level was also measured.

These tests consist of the most frequently mentioned types of reading material and require the testee to obtain the same kinds of information from the same manuals as job incumbents reported using in their work. Thus, they represent the most direct measure of actual job-specific reading task performance.

Each of the Job Reading Task Tests (JRTTs) constitutes a set of content-valid, job-specific reading tasks which can be used as a criterion-referenced measure of job reading task performance for that job. To facilitate comparison of job reading requirements, as defined by a JRTT, between different jobs and with other indices of job reading requirements, JRTT scores were scaled in terms of reading grade level, as measured by a standardized reading test. Thus, any job-specific JRTT score could be expressed as the performance of a soldier whose general reading comprehension was at some specified reading grade level, to use that common metric. In this fashion, and using the arbitrary but plausible decision rule that 70% of job incumbents should get 70% of the items correct on their JRTT, the reading requirement for Cooks fell at about the 7th-grade level, for Vehicle Repairmen at the 8th-grade level, and at the 12th-grade level for Supply Clerks.

Summary

These were our three main approaches to determining the reading level requirements of jobs. In each, we studied the relationship of general reading ability to a different criterion: measures of job proficiency, the structural properties of job reading materials, and the performance of empirically determined job reading tasks. These approaches agree in general in estimating substantially different reading requirements for different jobs, in a common ordering of the literacy requirements of the three jobs studies for all approaches, and in setting all job reading requirements at or above the 7th-grade reading level.

READING ABILITY OF PERSONNEL

Up to this point we have discussed research that has focused on the reading demands of Army jobs. The other side of the Army's literacy problem concerns the reading ability level of the personnel available to do the jobs.

Direct assessment of the current distribution of reading ability in the Army is not available because no reading test is standardly administered to all Army personnel. Throughout the several phases of this program, direct testing of reading comprehension has always yielded a substantial proportion of personnel of marginal reading ability. While this proportion fluctuates as a function of the quality of Army input, it is not likely to become trivially small. Most recent data as of March 1975 show that of 23,000 Mental Category III recruits screened on a standardized reading test, 17% failed to reach the 7th-grade reading level, and, of the 13,000 Category IV recruits screened, 43% were similarly reading at the 6th-grade level or lower. Subsequent retesting five weeks later in basic training reduced these percentages by half - which still leaves some 13% of Category III and IV input reading no higher than 6th-grade level at the time when they are about to enter job training.

An alternative to the direct assessment of the reading ability of Army personnel is provided in our study by the comfortingly consistent correlation ($r = 0.7$) between AFQT and reading grade level as measured by standardized reading tests. Since AFQT scores are available for all military personnel, the reading skill level of personnel in any job category can be reasonably estimated at any time from existing data banks.

Summary

To briefly summarize what has been learned about the Army's literacy problem: We have seen that (a) as determined by a variety of methods, the reading demands of Army jobs, even the less complex ones, far exceed the reading ability levels of many personnel, and (b) there is a consistent positive relationship between reading ability and job proficiency.

From this work we make the following conclusion: Although no single level of functional literacy can adequately represent the reading requirement of the range of MOSs studied, there appears to be a lower limit of 7th-grade reading level for functional literacy in the Army. Thus, remedial reading should be aimed at producing no less than 7th-grade reading ability and, optimally, should be targeted to the level of a man's MOS assignment. If elimination, retention, or promotion of career personnel is largely contingent upon paper-and-pencil job knowledge tests, formal procedures should be implemented to ensure that men have the opportunity to acquire both the job knowledge and the literacy skills required by the tests.

JOB FUNCTIONAL LITERACY TRAINING PROGRAM

Based on this research on the nature of its literacy problem, the Army in 1971 sponsored the HumRRO FLIT (Functional Literacy) project with the objective of developing a literacy training program designed to provide a level of functional literacy appropriate to present minimal job reading requirements and requiring no more than six weeks of training time. Given the absolute constraint of six weeks of training, we could see no reasonable prospects of increasing our adult students' general literacy competence to the point where it would transfer significantly to his job reading tasks. Accordingly, the FLIT objective was specified to be that of producing a student capable of using his job reading materials with the effectiveness of a man having a general reading ability of grade 7.0 or higher, as indexed by performance on a job reading task test. This led to the basic policy decision that all job reading training would be conducted using the concepts, content, and reading materials of the student's own job area. Parallel training curricula and materials were developed for each of six job clusters: Cook, Clerk, Communications, Combat, Mechanic, and Medic.

Entry

Entry to the program is governed by a series of screenings designed to eliminate most cases of testing artifact. All Mental Category III and IV input personnel are screened on a standardized reading test in the Reception Station during recruit processing. Those failing to reach the 6th-grade reading level are rescreened five weeks later in basic training. Those again scoring below the 6th-grade reading level are screened once more upon entering the reading training at the end of basic training by means of both a general and a job reading test. Only those students who have failed all three screenings are admitted to the job reading training program.

Program Overview

The functional literacy training program consists of three curriculum strands, each of which occupies about one-third of each training day. Strand I was designed to provide training in the application of existing general reading skills to job-specific Army job reading tasks. Strand II was designed to improve basic reading skills and job knowledge through using simplified versions of Army job reading materials. Strand III is devoted to general reading training and practice, as prescribed by the local Army school.

Strand I

Strand I training is designed to give the student drill and practice in applying his existing reading skills to the job reading tasks and the job reading materials which he will encounter in his entry level job

training and job performance. This Strand is a modular, self-paced, mastery-based program of job reading task training. Training is conducted in six modules, each addressed to one of the six fundamental job reading tasks identified as common and essential to MOS training and performance. These modules provide training in six job reading tasks: Using a Table of Contents, Using an Index, Using Tables and Graphs, Using the Body of the Manual, Following Procedural Direction, and Following Instruction in Filling Out Forms. Job-specific training is provided in using those Army manuals used in the student's own MOS cluster.

Entry to each module is determined by a module pre-test. Students meeting both time and accuracy criteria on a module pre-test advance to the next module; those failing the pre-test enter successive blocks of individualized instruction until they can pass a module post-test.

Strand II

In contrast, Strand II is a teacher-oriented program designed both to improve basic reading skills by language instruction and direct reading training, and to develop basic job knowledge in the student's own MOS field. In this curriculum, instruction is provided in decoding print to speech, in basic functional grammar and syntax, and in conceptualizing and structuring the meaning of a passage in linguistic, schematic, and pictorial representations. All instruction is keyed to sets of specially written, simplified reading passages presenting basic job vocabulary and job concepts in the student's job field.

IMPLEMENTATION

After a period of development and field trial in which all training was conducted by research personnel, this functional literacy training program has been implemented over the past year at all six Army Training Centers. The final phase of our R&D program consisted of evaluation of (1) the front-loaded functional reading training program in its operational setting, and (2) the feasibility of alternative delivery systems permitting the integration of the job reading training program with job skills training.

Let me summarize the findings of this study of the operational training program.

1. <u>The Current Need For Literacy Training in the Army</u>

In reading screening testing for this program in 1974 and 1975, 9% of Category III and 21% of Category IV recruits failed to reach the 6th-grade reading level on either of two administrations of a standardized reading test administered over the first six weeks of their Army service. Similarly, one-fourth of all input to a sample of entry level job training courses for drivers, mechanics, and supply clerks failed to reach the 6th-grade level of reading comprehension.

2. <u>Characteristics of Students Entering Reading Training</u>

Regional differences were substantial. Median amount of formal schooling ranged from 10 to 12 years, 29% to 57% of the students held a high school diploma or GED certificate, and 66% to 97% spoke English as their primary language.

3. <u>Program Effectiveness</u>

Pre- and post-module proficiency testing showed improved performance on all modules with substantial differences in the difficulty and effectiveness of various modules.

Summative evaluation showed high consistency of training effectiveness at the several installations. As measured by the *Job Reading Task Test*, mean reading grade level gain was 2.2 years, with 45% of the students reaching the 7th-grade reading level objective. On a *standardized reading test* measure, average gain was 1.0 reading grade level, with 29% of the students reaching the 7th-grade level. Retention testing eight weeks after completion of reading training showed that, for job reading, an initial gain of 2.4 RGL was reduced to a net gain of 1.9 RGL two months later. For the general literacy measure, the initial gain of 1.0 years was reduced to a net gain of 0.4 of a year.

4. <u>Feasibility of Integrating Job Reading and Job Skills Training</u>

As implemented, this literacy training program is front-loaded and adds six weeks to the training cycle when inserted between Basic Combat Training and entry level MOS training. Not only is this costly in terms of training time, but it also requires the learning of job concepts before the referents of those concepts have been experienced. Accordingly, studies were run on the feasibility of integrating job reading training and job skills training in one common program. When offered as either a part of, or an extension to, the regular job skills training day, the FLIT Strand I program obtained results directly comparable to those obtained in the front-loaded program, despite the curtailed time devoted to the integrated Strand I training. As in the case of the basic FLIT program, this integrated job reading training was a scheduled mandatory training activity for poor readers. When the FLIT Strand I program was offered on a voluntary unscheduled, walk-in basis at the post learning center, there were no takers. This lack of volunatry participation in the program may reflect the brief period of time the program was offered and the limited advertisement of the program. Future programs should expect to spend more effort in disseminating information about the program.

Summary

The Army-wide implementation of the front-loaded FLIT program, or the Advanced Individual Training Preparatory Training (AITPT) as it is now known, has just been completed in 1975. This implementation effort brings to a close this Army/HumRRO long-term program of literacy research and development.

CONTINUING ACTIVITIES

Currently, Army research and development in the literacy area is being conducted by the Army Research Institute. Ongoing is the development of a screening test for reading comprehension to replace the presently used USAFI test. Since the USAFI test is a military reprinting of a standardized reading test developed for use in the intermediate grades of the public school system, it requires academic knowledge and skills unrelated to Army reading needs. ARI is currently collecting pilot data on its new test of literal comprehension of Army-related reading materials.

Under consideration is a more extensive program of R&D aimed at the improvement of Army reading materials. The following initial major research activities are being considered:

1. The definition of relevant language variables and the development of computer programs to measure these variables.

2. The determination of those writing characteristics that need to be revised to substantially improve readability.

3. The development of computer programs to provide corrective feedback to writers.

COMMENTS ON THE PAPER BY CAYLOR

Richard Venezky
University of Wisconsin

Like others, Venezky based his comments not only on Caylor's paper, but also on other presentations and discussions. As preliminary observations, Venezky noted that he had gained a lot of respect from this Conference for the level of work going on in the Armed Services, and was surprised to discover that most of the military researchers were facing the same kinds of problems as those working in civilian settings with children. He noted that a fundamental difference between civilian- and military-based researchers was that the latter are almost forced to translate research into practice—to show results. University-based researchers, on the other hand, are rewarded not so much for applying what they learn to the solving of problems, but for publishing research and developing terminology which may in fact obfuscate how their research might be turned into practice.

Focusing on literacy problems within the Services, Venezky stressed four areas of concern. First, he expressed the idea that it is not literacy alone, but the *communication of information* which should be of concern to the Services. He noted that Caylor had reported a large influx of men with low levels of *language*, not just reading, skills; also, the Air Force is sponsoring work on the measurement of *auding* and reading skills. This suggests that the military's problem of communication extends to both listening and reading, not reading alone. Furthermore, the concern with readability research suggests that the Services are aware of *writing* as a problem. Hence, there is evidence that the Services must consider a total communications approach to their problems, and not focus just on reading.

From the communications perspective, Venezky suggested that, given a constant message one has to communicate, three factors are involved: the medium of communication (audio, visual); the arrangement and structure of the message (readability factor); and the competency of the receiver (language, literacy skills). Because research is currently going on in all of these areas in the Services, Venezky suggested that better coordination of these efforts is needed, because, for example, it is simply not cost beneficial to deal only with, say, literacy programs, when in fact the materials that people are expected to read are not very effective. Also, needs may change over time as costs change for print versus audio. Thus, there needs to be a coordinated research program looking at the total communications picture to better choose media for communications, to improve the structure and arrangement of messages, and to develop the required competencies of personnel.

A second, related area of concern expressed by Venezky was the need for a communications scale for various types of communications skills (literacy, writing, reading, speaking) and, in particular, a scale which

gets away from the use of grade scores. He felt that the latter scores are arbitrary, and in the educational setting of the military, are derogatory to the personnel. Rather than insulting personnel with such grade scores, Venezky suggested that the Services should talk in terms of the competencies necessary for the jobs to be done.

The third area of concern which Venezky discussed dealt with oral language. Many persons of low oral language skills, including many non-native speakers of English, are entering the Army, and it is necessary to teach them to comprehend the oral language, as well as the written language. With native English language speakers of low oral language skills, it seems necessary to improve these skills if their reading is to be much improved. Again, Venezky noted that the Air Force research on auding and reading indicated current awareness of this problem. He suggested that the services extend this concern with oral language, and focus not on teaching the grammar of English directly (the past perfect, the negative passive transformations, etc.) but rather, that a functional point of view be taken in which the functions of the language are stressed and trained.

Venezky's final comments concerned the need for the Military Services to explore alternatives to current training (literacy and technical) programs. Noting that the Services face the problem of limited time and funds for training, as well as high turnover of instructors, Venezky suggested the wider exploration of computer-managed or imparted self-study programs.

SUMMARY OF ONGOING RESEARCH AND DEVELOPMENT IN LITERACY TRAINING IN THE AIR FORCE

Steven D. Groff

Air Force Human Resources Laboratory

Before describing the current efforts in the Air Force's R&D program in the area of literacy training, it would be useful to note how the present training program is conducted and to gain some perspective on its effectiveness. All recruits are administered a locally developed fifteen minute reading test (designated RJS-1) upon their arrival at Lackland for basic military training. Those scoring below the sixth-grade level and all Mental Category IV personnel (those scoring below the 31st percentile on the Airmen Qualifying Examination), regardless of their performance on the RJS-1, are then administered the California Reading Achievement Test (CRAT). Those scoring below the sixth-grade level on the CRAT are assigned to the Reading Proficiency Unit for a period of up to eight weeks. In the proficiency unit the airmen's time is split between half a day of reading training and half a day of the usual basic military training.

The first week is spent in a conventional classroom setting working on work power and phonics. At the end of the week a diagnostic test is administered and the results used to guide the person's progress through the remaining seven weeks of the program. This latter portion of the course is self-paced and relies primarily on the SRA (Science Research Associates) Reading Series material, though other materials, such as McGraw-Hill and Readers' Digest, are available. Placement is automatically at the fifth-grade level and everyone is required to progress through the seventh-grade level materials before being eligible for early release from the Reading Proficiency Unit. Having achieved that point, the students are given the alternate form of the California Reading Achievement Test, and, if they score at the sixth-grade level or higher, they are channeled back into full-time basic military training. If they do not achieve that criterion level, they remain in the proficiency unit and are provided with additional training. At the end of the eight weeks, two courses of action are open for those who do not achieve the sixth-grade level. Either they are separated as having "limited potential" or, if the specific situation warrants, they are maintained in the unit for two additional weeks. This latter course is at the discretion of the unit commander and is used in roughly three percent of the cases.

Of the 439 entrants into the program in FY 74, 103 (23.5%) were discharged for reading problems, and an additional 68 (15.5%) were discharged for "other reasons." (The general attrition rate for basic military training in FY 74 was 7.7%.) The average length of time a recruit stayed in the Reading Proficiency Unit was 23 days, and the average grade level improvement was 1.8 (Nath, 1975).

Once out of basic training the airmen proceed either to technical school for anywhere from one to thirty weeks of training, or to a direct duty assignment. In terms of the reading improvement program, assignment to technical school merely delays the next step of the process. When a man with an AQE of less than 60 arrives at his first permanent duty station, he is given whatever reading test is locally available. If he scores below the ninth-grade reading level, he is automatically enrolled in the reading program offered by the base education office, and supposedly, continues until he achieves a ninth-grade reading level. Even though a person may initially test above this criterion, he may be required to enroll in the reading program if he has trouble completing upgrade training and/or his supervisor somehow determines that his problem is one of reading. These base-level (local) programs vary rather dramatically since the requirement is only for their existence. As a result, programs are established on the basis of what is available in the community; this may be a transplanted public high school/GED course or it may be a course provided through a special contractual agreement with a university or other institution.

There is no easy means of tracking all the existing reading programs, but it seems reasonable to make the generalization that the typical approach has a very traditional orientation. Such a program is conducted at Lowry AFB, Colorado, through an arrangement with the Denver Public School System. The program is essentially indefinite in length since anyone who desires to attend the twice-a-week class meetings may do so for as long as he likes. Roughly an hour of the class is spent in lecture/participation format covering phonics, work attack skills, and the like. During the second hour the students work with SRA materials and receive individual tutoring as required. Periodically, testing is done to *informally* track progress but no permanent records are maintained.

The extent to which these programs attempt innovation is quite limited. The Education Services Office at Kirkland AFB, (Mr. Illes, personnel communication), for example, has instituted an eight-week program under contract with the University of New Mexico. Of twelve hours of instruction per week, four are in the area of study skills, involving the use of dictionaries and orientation trips to libraries. Two hours are spent in group counseling sessions (a rather unique feature) and the remainder of the time is spent in other reading activities using Air Force Career Development Course materials. During the first six weeks, heavy emphasis is placed on phonics because of an apparent need for such training. The final two weeks are spent building on this background by having the group read aloud plays which are then discussed and which serve as a vehicle for further skill development.

IS THERE AN AIR FORCE LITERACY PROBLEM?

Concern has been expressed at the 1972 and 1973 World Wide On-the-Job Training Conferences by supervisors and training managers that reading problems are proving detrimental to the conduct of training. Their

position appears to be supported by some recent research. As a result of a survey by Mockovak (1974a) it was found that 5,774 airmen participated in base-level reading programs between 1 April 1972 and 1 April 1973 and that the most frequently cited reason for enrollment was the inability to read and/or comprehend Career Development Course material (that is, those self-study materials required for skill upgrading and promotion). In addition, base education officers reported that, on the average, fifty percent of those enrolled in their reading programs were high school graduates, and that eighty-four percent of those enrolled represented only ten career fields. Thus, there is the further indication that the more highly educated people within the Air Force manpower pool are not all adequate readers. In absolute numbers, the existence of a "reading problem" appears to be a legitimate concern, particularly if one realizes that the turnover of personnel is a long-term process and that enlistment criteria are likely to fluctuate downward from the January 1975 level.

Mockovak (1974b) studied the literacy skills of AF personnel and the reading requirements imposed by AF training and job materials. Mockovak estimated the reading ability of three- and five-skill level personnel using the Madden and Tupes (1966) regression equation for AQE (Airman Qualifying Examination) scores, and also estimated reading requirements using the FORCAST (Caylor, et al., 1973) readability measure. Identifying the reading requirement level as that Reading Grade Level (RGL) below which seventy-five percent of the material fell, he compared this requirement level to the mean RGL of AF personnel by career ladder. This calculation resulted in a "literacy gap." Of the 56 ladders he studied, Mockovak found forty-three to have a negative gap (i.e., a reading requirement of materials exceeding the mean reading ability of those who were expected to use them). Of these forty-three, twenty-nine had a negative literacy gap greater than one, seventeen greater than two, and four greater than three; thus, thirty percent of Mockovak's sample had a negative literacy gap of greater than two reading grade levels. These results tend to confirm some of the increasingly frequent comments by training managers that reading problems are having a detrimental impact on the on-the-job training program. However, it is useful to note that these data do not indicate that a specific problem exists or what the operational consequences of the subjectively defined "literacy gap" are.

Other data which provide an important perspective on the scope of the problem, and suggest that it may be underestimated, were compiled by DeGuelle (1974). He compared performance of airmen scoring at the ninth-grade (the ceiling) on the USAFI III reading test with their performance on the SRA Diagnostic Reading Test, Survey Section, Upper Level, (grades seven through college freshman year) Form A. (Within some AF commands the USAFI is widely used). Using ninth-grade norms on the vocabulary portion of the SRA, 24% of the testees were identified as having reading deficiencies; on the total comprehension test, 22% were identified as deficient. Using twelfth-grade norms, 68% of those tested were deficient in vocabulary, and 70% in total comprehsnsion. DeGuelle's operational

definition of "deficiency" on the SRA was the achievement of scores at or lower than the 30th percentile.

This suggests that estimates of the number of persons with generally inadequate reading skills are likely to be low. To what extent the same can be said of job specific reading skills is unknown.

A final consideration before progressing to the current AF R&D in literacy training is the recent change in enlistment standards which will surely have some effect on AF reading training, particularly on the operation of the Reading Proficiency Unit described earlier. In January 1975 the enlistment standard was raised to a General Aptitude Index (GAI) score of 45 with a composite of 170. Shortly thereafter the Air Training Command (ATC) initiated a 90-day data collection effort at Basic Military Training at Lackland AFB. A forty-percent random sample of recruits was administered the California Reading Achievement Test. A total of 5154 non-prior service personnel were tested. Forty-three percent of those tested scored at the thirteenth-grade level and only 3.7% scored below the ninth grade. (The officially recognized standard for reading ability is the ninth grade.) The mean reading grade level was 12.1 and the standard deviation was 1.6. Assuming this enlistment standard is maintained, no one with a GAI less than 45 should be entering the system, the expected result being that there will be essentially no one below the ninth-grade standards. This is an important consideration in planning the direction of AF research. At the same time, though, historical trends suggest that entrance standards are likely to fluctuate in the future (Dailey, 1951; Lecznar, 1962; Vitola, Valentine & Tupes, 1967); with the current high standards, that fluctuation will be downward.

CURRENT LITERACY R&D IN THE AIR FORCE

Research and development in the AF has only recently begun to consider in a systematic fashion the area of literacy training and two major efforts are currently ongoing. The first of these calls for the development of a literacy assessment battery and is based on Sticht's auding-reading model (Sticht, et al., 1974).

The rationale for the literacy assessment battery lies in the basic processes involved in the acquisition of language. As a person moves from an initially pre-linguistic stage and interacts with his environment, he begins to develop auding and speaking skills. These aspects of languaging continue to develop as the person enlarges his conceptual base and begins to make associations between written and spoken language. As the person progresses through the early years of school, he begins to acquire those skills involved in reading and writing; up until about the eighth grade (that is, through stage 3 of Sticht's model) a person's oracy skills generally surpass the "literacy" skills of reading and writing. At about that point the person has generally acquired the prerequisite decoding skills which permit the oracy-literacy gap to narrow, after which the skills then develop simultaneously.

The extent to which cognitive content and processes have developed is a constraint upon the acquisition of further oracy and literacy skills. Also, within this framework, the development of literacy skills is limited at any given time by the current level of oracy skills as well as the degree of mastery of decoding skills. The corollary to this is that, practically, one can only hope to improve literacy skills up to the level to which oracy skills have been developed.

It is upon this proposition that the Literacy Skills Assessment Battery is being developed. It is intended primarily for the screening of persons with a reading level of eighth grade and below, and currently is composed of three sections: paragraph comprehension (with auding and reading components), vocabulary, and tracking. The first section will yield a "potential" score indicating the discrepancy between the person's auding and reading skills. The latter two sections fulfill an essentially diagnostic function (for decoding, vocabulary, and automaticity) for problems that surface in the first section. The Literacy Assessment Battery is being designed to overcome the shortcomings of currently available types of reading potential tests (e.g., the Diagnostic Reading Scales and the Durrell Listening-Reading Series) and its advantages over these tests are that 1) it is adult oriented, with the passages drawing upon experiences familiar to young adults; 2) the questions are passage dependent and open ended; and 3) the vocabulary section is keyed to the comprehension passages.

In light of the previously noted inadequacy of some general reading tests, it is hoped that this new Literacy Assessment Battery will prove useful in conventional reading development programs, as well as in the job-oriented type program, which constitutes a second area of current Air Force research and development.

In view of Mockovak's literacy gap data and apparent problems with the Career Development Course materials, the Air Force Human Resources Laboratory (through a contract with the Human Resources Research Organization, HumRRO) is now developing a prototype job-oriented reading program for the Air Force. The approach being taken is to identify job reading tasks that are commonly encountered on the job and in the self-study job training materials, and then to train those reading skills that are required to accomplish the tasks. A similar approach has been used to develop a functional job reading program for the Army (Sticht, 1975b); however, the nature of the Army's problem is rather different than that of the Air Force, Evidence for this derives from the fact that substantial Air Force technical training already involves procedures on such tasks as how to fill out forms and locate information in manuals: tasks which are taught in the Army's FLIT (Functional Literacy Training) program. Also, data has been collected by Dallman (personal communication) on HumRRO's Mechanic Job Reading Task Test. Dallman administered the test to thirteen airmen *prior* to their entrance into the vehicle maintenance course at Chanute AFB, Illinois, and found that seven persons scored in the highest possible category and the range of scores was quite narrow, 11th-14th grade levels. Thus, there is the indication that attention must be focused on the higher level comprehension skills.

It is anticipated that much of the materials developed as part of this effort will deal with teaching airmen to perform transformations on textual materials, such as transforming connected prose into flow charts, classification tables, and pictures, in such a way as to more completely understand the materials. The working materials will be extracted (and often modified) from Career Development Courses and Technical Orders. Materials will be developed for use within one or two clusters of two or more career fields. If two clusters are eventually produced, one will represent a maintenance (aircraft maintenance and repair) and the other a non-maintenance (administration/personnel) area. If the prototype program is successful, it is likely that additional modules will be produced for at least those career fields that are experiencing the greatest problem.

Specially developed job-oriented reading programs such as anticipated here are one approach to AF reading problems and the typical lack of motivation that they inspire. It is conceivable that such a program could function as a vital part of a total career education program. A Reading Proficiency Unit located at Lackland AFB might function in a manner similar to the present mode for those recruits of low *basic* reading skills; the job-oriented program could exist as a next step and concentrate on improving the conceptual behaviors of the airman within a job context; and finally, a reading/GED-type program could provide access to that higher level of skill which is only attainable over rather lengthy periods.

Closely related to the job-oriented reading program's emphasis on conceptual activities is the consideration of ways in which these activities may be made more efficient. The study to be described here has been completed relatively recently by Dansereau, Long, McDonald, Actkinson, Ellis, Collins, Williams, and Evans (1975).

The rationale for this study derives from findings that instructional techniques and sequences of information presentation have very little, if any, impact on subsequent performance (Dansereau, Evans, Wright, Long, & Actkinson, 1974; Dansereau, Evans, Actkinson, & Long, 1974). Also there is the suggestion by Dansereau, Long, McDonald, and Actkinson (1975) that the effective learning strategies which facilitate a flexible approach toward the learning task may be a prime source of student problems. Difficulties in comprehension and retention were viewed as stemming from two sources: "first many students tend to receive information passively and consequently do not actively integrate it into their existing cognitive structures...(and) second, many students apparently do not attempt to produce multiple memory representations (that is, encodings) of the same material in order to enhance retrieval" (Dansereau, Long, McDonald, Actkinson, Ellis, et al, 1975, p. 12).

Following this reasoning, instructional materials were prepared for three techniques that, when used by the student, were expected to aid comprehension and retention: question and answer, paraphrase, and visual

imagery (see Dansereau, Actkinson, Long & McDonald, 1974). Each package also contained a section on information retrieval. Each of three groups of subjects was presented with a different technique package and trained and tested in a series of four sessions totalling about seven and one-half hours.

Analysis of the results shows no performance differences on the immediate post-test, which consisted of 80 true-false, multiple, fill-in-the-blank, and short answer items. However, the paraphrase and imagery groups performed significantly better than the control group when given an "essay type" test five days after reading the assessment passages. Based on the findings and recommendations of this effort, it is anticipated that this area of study will be pursued using job-related materials both within the conventional and job-oriented training programs.

In summary, then, the current Air Force research program in literacy training involves the identification of those persons who would most likely benefit from further training and the development of a relatively short-term program to train techniques which facilitate the comprehension of job and training materials. Both efforts described with respect to this latter category rely heavily on the manipulation of the forms of informational representation; such manipulation, theoretically, will aid the integration of new material with existing cognitive content. A definitive evaluation of the success of this approach awaits both the development of the job-oriented reading program and, in the longer term, a valid method for evaluating the operational consequences of the "literacy gap."

COMMENTS ON THE PAPER BY GROFF

Samuel Gibbon

Children's Television Workshop

Gibbon based his comments not only on Groff's paper, but also on the other papers and discussion at the Conference. He expressed his point of view that the Armed Services appear to be proceeding with enlightened programs of research and development towards closing the gap between reading skills of personnel and reading demands of jobs. However, these Services at times have had the option of screening out many of the people who do not measure up to their standards. The choice of selection standards and the resulting implications for the marginal citizen may require some concern by all of us about the role of the Armed Services in society. For instance, the view had been expressed in some discussions that the criterion for determining the effectiveness of and need for literacy training in the Army was whether or not it improved the combat readiness of the Army. Gibbon suggested another basis on which literacy training in the Army might be undertaken: that of the Army as a place which provides opportunity for the marginal citizen, and that is a considerable social function for that Armed Service to perform. He speculated that there will be a good many people who will enter the Army not knowing how to read well, and offered his conviction that people who do not know how to read want very much to learn how! They may conceal that desire, but they indeed want to. Thus, because the Army has been successful in teaching many other skills, it could, if it chose, teach these people who desperately want to learn to read but have not, how to read. Gibbon suggested that there are many ways in which reading training could be done, but pointed to the need for a values orientation such that developing the ability to read a simple sentence might have the same importance in basic training that a shoeshine does!

Pointing to the success of television and other electronic media in providing private ways of learning for people whose shame at their ignorance is extreme, Gibbon suggested that use of such media in literacy training might be explored in all the Services. Finally, returning to Groff's paper, Gibbon commented on the observations in that paper which indicated that the operational consequences of a literacy "gap" have not been identified, and expressed his belief that discovering those consequences might contribute to the program of ongoing research and development.

ONGOING RESEARCH AND DEVELOPMENT IN READABILITY AND USEABILITY OF TECHNICAL WRITING IN THE ARMED SERVICES

Preceding page blank

INTRODUCTION

The preceeding section dealt with one main approach to literacy problems in the Armed Services: R&D to produce better readers through literacy training. This section deals with the second main approach: R&D to produce better reading materials.

The section begins with Tom Curran's review of readability research in the Navy. He begins with a discussion of the terms "readability" and "comprehensibility" and of the interaction of readability, comprehensibility, and useability with each other and with various characteristics of the readers.

Curran discusses both the prediction of readability and the production of readable writing. The prediction of readability is attempted through application of a readability formula. Curran describes the nature and application of five readability formulas and considers the special problem involved in assessing the readability of technical material and the tradeoff between the difficulty of computation of multivariable formulas over that of one- or two-variable formulas and the resulting increase in predictive power. In addition, he discusses non-statistical methods of measuring difficulty, especially the cloze test measure of comprehensibility. In describing the process of producing readable writing (including the role of prediction), Curran discusses the role of style guide and military specifications. He describes several automatic and computerized systems which can be used to aid the writer by automating the process of prediction.

The focus shifts to comprehensibility as Art Siegel reviews Air Force research and development in that area. Such R&D includes attempts to use auditory supplementation to increase comprehensibility and the development of new comprehensibility measurement concepts which focus on the cognitive and language processing load which a text places on the reader. Siegel concludes with a description of a comprehensibility measurement computer program.

Dick Kern reports on the Army's R&D efforts not only on the reading difficulty of Army literature, but also on its content and orientation and its useability. Kern describes studies which investigated the relationships between the reading ability of personnel and the reading difficulty of the manuals they were expected to use, their use of manuals on the job, and the quality of their performance on job sample tests. He reports on research which indicates that Army training literature in the past has not always been oriented toward its users, both in content and in orientation.

The Army is attempting to develop more user-oriented literature, as indicated by a user-oriented approach in a guidebook for Army writers and by the user-oriented publications of CATB. Related to the question of user orientation of a manual is the question of its useability. As Kern reports, most past evaluation of manual effectiveness has been in the context of entire training programs. Kern suggests that future research on the useability of printed material focus on the identification of characteristics of printed material which influence useability followed by the development of methods to assess useability of materials.

Much of the Air Force R&D on the useability of job reading materials has been connected with the development and evaluation of job performance aids which are various printed materials designed to enhance on-the-job performance, especially of maintenance and troubleshooting tasks. In reporting on those efforts, Robert Johnson discusses specifications for the development of useful job performance aids (i.e. the steps to take to ensure the development of useable data) and the resulting characteristics of the completed JPA s which contribute to their useability.

Job performance aids have also been a part of the Navy's R&D on the useability of technical manuals. Bill Muller, in his review of current Navy research on useability, graphically shows the increasing magnitude of the production of technical manuals. He discusses the key elements in the Navy's technical documentation system and some of the problems of the system with regard to the development of useable technical documentation. Muller concludes his report with a review of four ongoing efforts which attempt to provide assistance for the technical writer.

READABILITY RESEARCH IN THE NAVY

Thomas E. Curran

Navy Personnel Research and Development Center

A great deal of concern has recently been voiced with regard to the quality of instructional and maintenance manuals in the Navy. Men, both on the job and in the classroom, have tended to have problems using manuals due to the difficulty of both text and graphics, inadequacies in content, lag time in updating, etc. The thrust of this paper is with the first of these deficiencies, but its concentration differs from that often found in work on readability. Readability, *per se*, is not a difficult problem if one assumes as its operational definition a "readability formula score." As measured by readability formulas, almost *anyone* can write readably. One simply uses short words and short sentences. Given these, the "readability" of the material will be at a high ("easy") level when indexed by the most common formulas. Obviously, this is a simplistic view, particularly for technical writing. Technical writers (who encompass nearly all writers preparing Navy materials) cannot always obey the dictum to use short words. What "easy" word can be substituted for "oscilloscope," for example? But even for these writers, readability need not be a severe problem. One must merely modify the above prescription to say "use *familiar* words and short sentences" and modify the readability assessment procedure so that *long* familiar words do not inflate the reading grade level (RGL) of the material. A major part of the R&D effort in the field of readability should therefore be directed at identifying long words which are known to be familiar to the intended audience and effectively transmitting this information to the writer. This topic will be discussed in other contexts at a later point.

There is a second characteristic of writing that deserves special attention—comprehensibility. There seems to be some misapprehension in the literature to the effect that high readability *ensures* comprehensibility. Only if readability is taken to include good "style," smooth flow of ideas, avoidance of complex sentence structures, etc. is this the case. Note that this represents a rather drastic departure from the operational definition of readability suggested above, and, in fact, closely resembles the accepted definition of *comprehensibility*. But the position is taken here that they are clearly two different concepts. At this point in time, readability formulas cannot adequately assess characteristics such as style and sentence structure. It is true that readability and comprehensibility often go hand in hand. The nursery rhyme "Mary Had A Little Lamb" is written at a very low level of difficulty, and because of the structure of its sentences, its simple words and its easy flowing style, it is also quite comprehensible. But examine the

other extreme (unlikely as it would be coming from any rational author). If the words in the rhyme were "scrambled," with the length of sentences left intact, its readability score according to formula would be identical to that of the original version. Yet its comprehensibility would be reduced enormously. In between these two extremes fall countless examples of writing which vary along each of the two dimensions. In short, comprehensibility, like readability, is a property of the written material, and when the work is done carefully the two are at least moderately correlated. But it would seem prudent to continuously bear in mind that the former cannot be directly assessed by means of existing formulas.

To further complicate the issue, there is really no point in speaking of *either* readability or comprehensibility unless the intended audience is taken into account. Two characteristics of the audience (other than general reading ability) are of prime importance: the background knowledge possessed, and the degree of motivation involved. The adult reader faced with the scrambled version of "Mary Had A Little Lamb" might comprehend it reasonably well because of repeated exposure to it in the past, and might find it to be a stimulating (i.e., motivating) puzzle-solving experience. The same strings of words seen by a child for the first time would probably be incomprehensible, even though the individual words themselves could be decoded.

Lest one be mislead by the above statements, it should be made clear that readability is NOT a trivial issue. It is a *necessary*, but not a *sufficient* condition for high-quality manuals. All other factors being equal, COMPREHENSIBILITY requires high READABILITY, and one step removed, USEABILITY requires high COMPREHENSIBILITY. If a manual is not readable, the student or technician will not even attempt to use it once he has been stymied by its difficulty at the outset. He probably will avoid it from that time on, just as would anyone presented with a tome replete with long words which he cannot decode. But to be readable is not enough. If the manual is not also comprehensible, the man possibly will not be able to use it despite his need for the information it contains. And the problem again is not even that simple. Material can be virtually unreadable according to a formula score and yet one who is vitally interested in the topic and/or needs the information regardless of the effort required to obtain it may slave his way through it despite the writer's opposition.

The interactions among readability, comprehensibility, and useability (all properties of the *material*) and reading ability, comprehension ability, motivation, and background experience/knowledge (all properties of the *user*) should provide the overall model for any R&D effort in improving Navy manuals.

In the third quarter of Fiscal Year 1975, the Navy Personnel Research and Development Center (NPRDC) was tasked to investigate the area of the readability and comprehensibility of technical manuals (TMs). An

extensive survey of the literature was undertaken to determine the state-of-the-art in this area and report findings (Curran, 1976). The conclusions of this author were that the area of readability and comprehensibility has been covered in breadth but not in depth. A dilemma appears to have evolved. Formulas can be applied efficiently—automation has become the rule—but it is possible that only judgments of the material by the readers can provide an index of its comprehensibility.

PREDICTION OF READABILITY

This author makes the distinction between prediction of readability and production of readable writing as put forth by Klare (1975). This section deals with the first of these problems while the following section deals with the latter and the interaction between the two.

Readability Formulas: General

The most common method for predicting the difficulty of a given piece of material is the readability "formula." A great many of these have been developed, with the majority using as variables some index of sentence difficulty and individual word difficulty. Reviews of these formulas up to 1960 can be found in Klare (1963); Klare (1974-1975) reviews those developed after 1960 and modifications to earlier ones. Reference in this paper will be made to only five such formulas: the FORCAST formula, the RIDE scale, the Fog Count, the Flesch Reading Ease (RE) formula, and the Dale-Chall formula.

In general, the prediction of readability by formula involves "counts" of various components of written material and, using passages of known difficulty as criteria, computing a regression equation on the counted characteristics. Commonly, the criterion passages are developed using the "cloze" technique. This procedure, which is of concern throughout this paper, involves extracting verbatim passages from the material and deleting every nth word, replacing each with blanks of standard length. Normally, every 5th word is deleted, resulting in five versions of the test for each passage (deletion of words 1, 6, 11, ... n, up to 5, 10, 15 ... n). Thus, each word in the passage is deleted once across the five versions. Subjects of known reading ability are asked to fill in as many of the deleted words as they can, without having first read the intact passage. The reading level at which 50% of subjects can fill in approximately 40% of the missing words is usually taken as the reading grade level (RGL) of that passage.[1] This "scaled" RGL is then used as the criterion for regression of the counted variables.

---

[1] RGL is comparable to, but not equivalent to, school grade; it is a somewhat arbitrary level at which a particular grade student *should* be able to read with satisfactory comprehension. 40% correct on the cloze test has been shown to be approximately equivalent to a 75% score on a multiple-choice test on the material.

As an example, Bormuth (1969) used this procedure for deriving a number of readability formulas. He examined, in an extensive correlational study, approximately 170 different variables—such as vocabulary, syntactic complexity, and parts of speech—and their relationship to the difficulty of written material. He first determined the correlation of each of these variables with the difficulty of 330 100-word passages as scaled with the cloze technique. He then entered the variables into a regression equation to determine which of them in combination best predicted the difficulty of the passages. The outcome of this study was a series of formulas, each designed for a different purpose, incorporating the "best" of the variables for predicting the difficulty of other writing.

The Flesch "Reading Ease" Formula

Probably the most widely used (and most consistently powerful) of the readability formulas is that developed by Rudolf Flesch (1948). This formula - termed the Reading Ease (RE) formula - uses as variables the number of words per sentence and the number of syllables per 100 words of text. The original RE formula is presented in Table 1a. This formula was developed in much the same manner as described above, except that the criterion was a set of standardized passages rather than passages normed for the specific purpose by the cloze technique. The formula indexes the difficulty of material on a scale from 0 (practically unreadable) to 100 (extremely easy). Using this index, the RGL can be determined from a conversion table. The RE formula was developed and validated on children and civilian adults, with reading material appropriate to these samples, and its usefulness for assessing the difficulty of military technical or instructional writing is therefore suspect. In order to overcome this problem, Kincaid, Fishburne, Rogers, and Chissom (1975) recalculated the formula using Navy enlisted personnel reading Navy job-relevant literature.[2] In the process of this recalculation, the RE formula was also revised to produce RGL directly without having to take the extra step of converting "reading ease" to RGL via a conversion table. The recalculated version of the RE formula is shown in Table 1b. The RE formula has the advantage of being relatively simple to compute manually and is adaptable to automation as well. Klare, Rowe, St. John, and Stolurow (1969), among others, have developed a computer program which provides the RE index based on the original Flesch formula. It remains a fairly simple task to revise the program for the recalculated version of the formula.

---

[2] These authors also recalculated two other formulas - the Automated Readability Index and the Fog Count—which will be addressed at a later point.

TABLE 1a

ORIGINAL FLESCH READING EASE FORMULA

$$\text{READING EASE} = 206.835 - 1.015\,(\text{words/sentence}) - .846\,(\text{syllables/100 words})$$

TABLE 1b

RECALCULATED FLESCH READING EASE FORMULA

$$\text{GRADE LEVEL} = .39\,(\text{words/sentence}) + 11.80\,(\text{syllables/word}) - 15.59$$

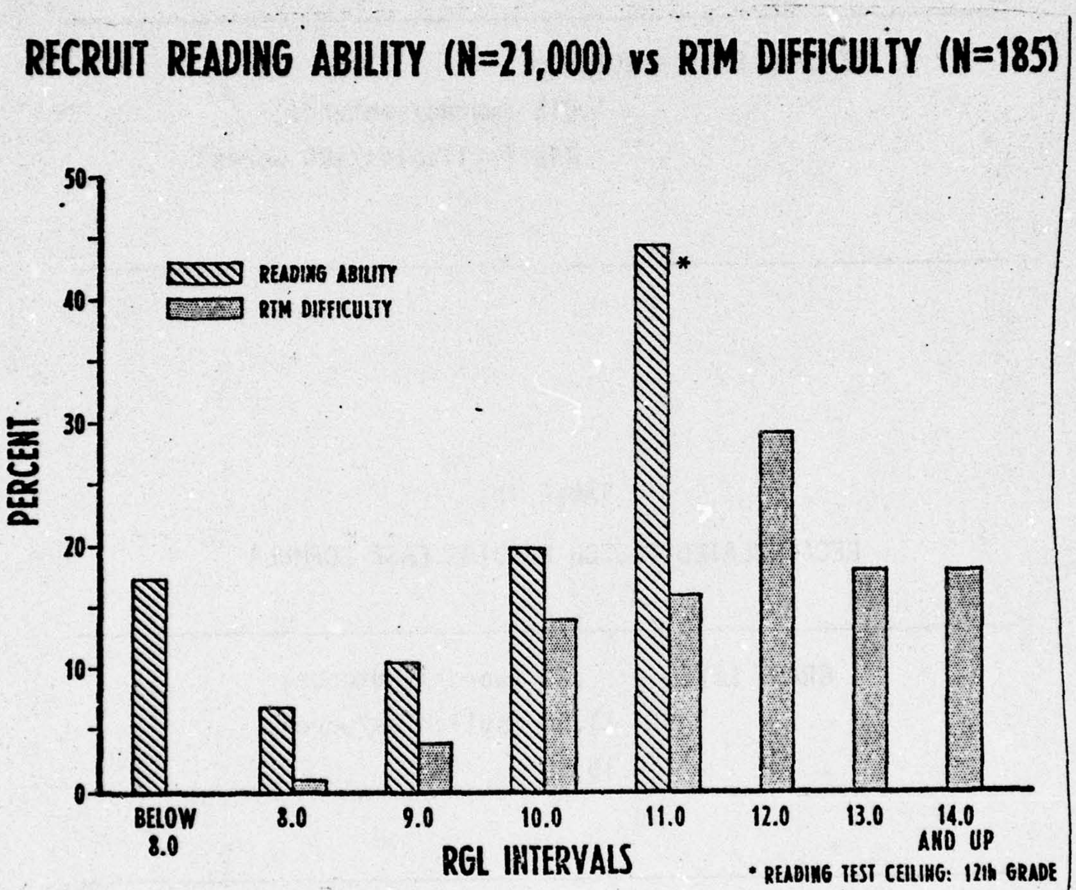


Figure 1. Comparison of Recruit Reading Abilities and Training Manual Difficulty.

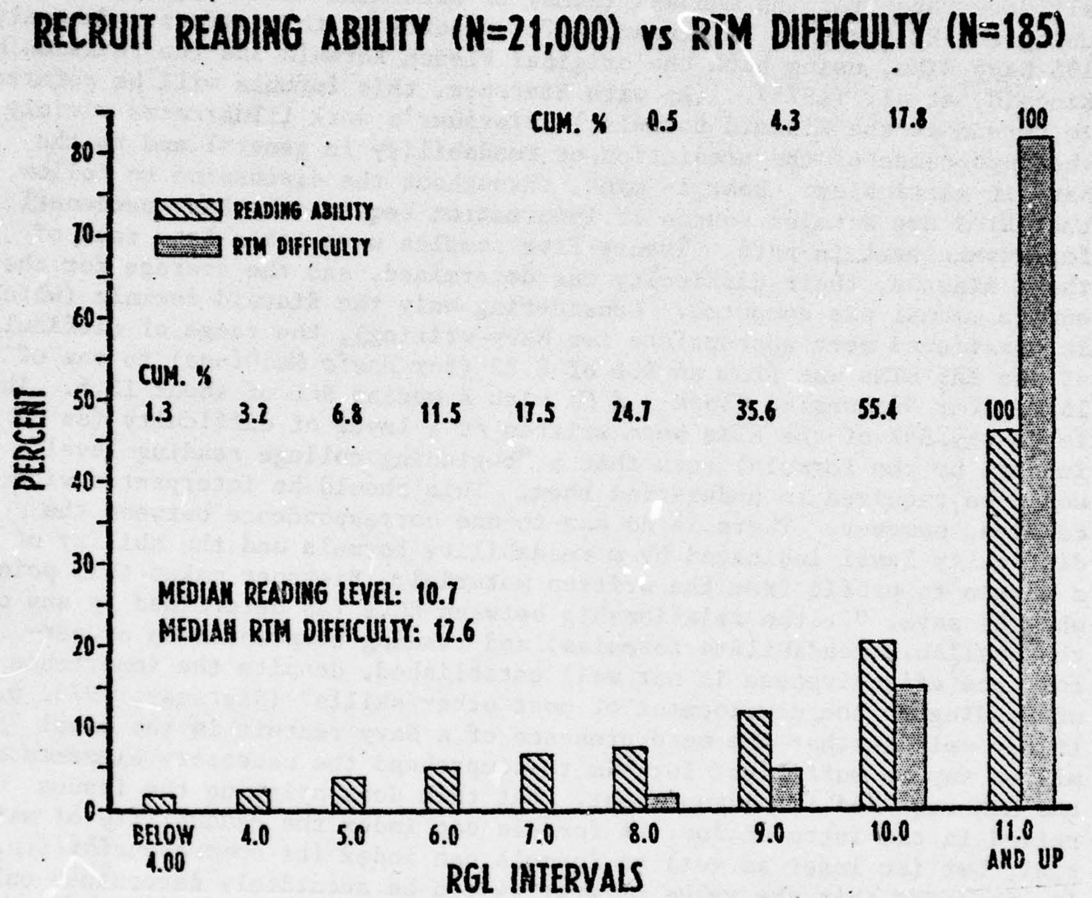


Figure 2. Comparison of Recruit Reading Abilities and Training Manual Difficulty

Application of the Reading Ease Formula

Based on its consistent validity when compared with other formulas and its relative ease of manual computation, the Chief of Naval Education and Training Support "Readability Working Group" conducted an analysis of all Navy rate training manuals (RTMs) to determine their difficulty using the RE formula. Biersner (1975) reports on this effort. He analyzed 195 Navy RTMs, using both the original Flesch formula and the revision by Kincaid, et al. (1975). (As with Biersner, this formula will be referred to herein as the Kincaid formula.) Biersner's work illustrates vividly the importance of the prediction of readability in general and to the Navy in particular. Bear in mind, throughout the discussion to follow, that RTMs are a major source of information required by Navy personnel for advancement in rate. Twenty-five samples were taken from each of these manuals, their difficulty was determined, and the average for the entire manual was computed. Considering only the Kincaid formula (which is considered more appropriate for Navy writing), the range of difficulty of the 185 RTMs was from an RGL of 8.82 (for *Basic Machines*) to one of 16.26 (for *Disbursing Clerk 1 & C*) with a median RGL of about 12.6. That is to say, 50% of the RTMs were written at a level of difficulty (as indexed by the formula) such that a "beginning college reading level" would be required to understand them. This should be interpreted with caution, however. There is no one-to-one correspondence between the difficulty level indicated by a readability formula and the ability of a person to profit from the written material. Biersner makes this point when he says, "...the relationship between RGLs (as determined by any of the available readability formulas) and reading comprehension or performance effectiveness is not well established, despite the importance of reading to the development of most other skills" (Biersner, 1975, p.7). It may well be that the mere presence of a Navy recruit in the naval milieu may be sufficient for him to comprehend the necessary elements of the RTM required for advancement. But this does point up the issues raised in the introduction. A formula can index the *readability* of material, but (at least as yet) no formula can index its *comprehensibility*. It is clear that the value of writing can be accurately determined only from within the framework of an overall model - the readability *and* the state of the user in terms of motivation, background knowledge and experience, etc. But at face value, assuming all other factors to be equal, the comparison between the RGL of the manuals as reported by Biersner and the reading abilities of recent Navy recruits (see Duffy & Nugent, 1975) shows a quite apparent "man-manual mismatch." This comparison is presented in Figures 1 and 2. Figure 1 indicates the full range of the RGLs of the RTMs from the one manual at the 8th-grade RGL to those at the 14th and above. Reading abilities of recruits (as measured by the Gates-MacGinitie test) are shown for the same intervals. The fact that no reading abilities are shown for levels above 12th grade is an artifact, due simply to the fact that the ceiling for this test is at the 12.0 level. Figure 2 presents the same data in a different

perspective. Here, reading abilities are indicated by percentages from grade 4.0 (and below) to grade 11.0 (and above). All RTMs whose difficulty exceeded 11.0 according to formula score are bulked into the latter interval. Noting the cumulative percentages (again at face value), it would be possible that approximately 24% of these recruits could read only *one* RTM—the single manual in the sample below the 9.0 level. To interpret further, approximately 82% of the RTMs are written at the 11.0 level or above; only 45% of the recruits were tested as reading at the 11.0 level or above. And one further illustration of the mismatch (not indicated in the figures) deals with the manual *Basic Military Requirements*. The RGL of this manual is 10.85. Knowledge of its contents are required for advancement to *E-3*. Yet its difficulty level is approximately the same as the *median* reading level of the over 21,000 recruits! It is possible that 50% of the entire sample would not be able to read this manual. Clearly, our ability to predict the difficulty level of written materials is of immense value. The data discussed above gives us every indication that certain RTMs should probably be revised downward in difficulty level unless reading abilities can be increased.

In addition to the striking findings reported above, another feature of Biersner's work deserves mention. As noted earlier, the computation of the reading ease formula involves a count of syllables, words, and sentences. While this can be accomplished relatively simply with nothing more than a paper and pencil, Biersner reports on a device which greatly facilitated the giant task of his analysis. Biersner directed the development of an electro-mechanical counting device which would permit more rapid and reliable gathering of the data. This device took the form of a "stylus" which, when pressed to a surface, trips a microswitch, which in turn activates a counter. A relatively unskilled person using this device can therefore press the stylus to the working copy at the end of a syllable, word, or sentence, and that variable would be indexed by the counter. Biersner reports that "...it made data collection over 30 percent faster, while maintaining high reliabilities" (Biersner, 1975, p. 17).

Technical Terms and Measurement of Difficulty

A major problem with most readability formulas (as alluded to above) is that they were developed and validated using either school children or civilian adults. There is good reason to believe that the abilities of a child reading at the 6th-grade level (for example) and of a Navy man also reading at the 6th-grade level are quite different. Again, the military environment itself may be expected to account for perhaps great differences in the words with which the person is familiar. This problem is even more acute when one is concerned with technical writing. Since word difficulty is usually indexed in terms of the length of the word, the technical terms encountered on the job will tend to inflate readability formula scores. Yet is is probably nonsense to assume that an

electronic technician would be unfamiliar with the word "oscilloscope." Traditional formulas have not addressed this problem directly, and yet they must. This is central to the ongoing theme of this paper. The fact that a readability formula indexes a piece of writing as "easy" or "difficult" does not ensure that it will appear that way to the intended user. Two questions seem to be at issue here: to what degree *do* technical terms inflate the analysis (since they seldom exceed about 15% of the total number of words), and, if indeed they are a problem, by what means can the difficulty of technical writing be realistically indexed? The first of these questions is a basic empirical one, and should be answered before proceeding with R&D involving the second question. Additional formulas, and measurement considerations not involving formulas, will be discussed below, taking these R&D priorities into consideration.

The FORCAST and RIDE Formulas

Caylor, Sticht, Fox, and Ford (1973) made an attempt to allay the problems of using formulas derived from civilian samples in determining the difficulty of military writing. They examined a total of 15 structural variables in Army training literature, including words per sentence, number of independent clauses, number of one-syllable words, and total number of syllables. They were also dissatisfied with the traditional readability formulas for use with military writing:

> The fact that formulas have validity coefficients of about .70 for predicting the performance of *school children* on reading comprehension tests indicates that they account for roughly 50% of the variability in reading performance of *children*. It is likely that they may account for less variability in *adult* performance, especially since material containing large numbers of technical terms would increase the estimate of difficulty made by the readability formulas. (Caylor, et al., 1973, p. 6.)

They therefore examined job-relevant materials from seven military occupational specialties (MOSs) in the Army, using scaled RGLs of this material (based on the cloze test) as a criterion, and computed regression equations with the 15 variables. They found that the number of one-syllable words alone correlated .86 with the criterion cloze scores, and that the addition of either one or two additional variables to the equation produced no significant increase in this correlation. They therefore included only this variable - number of one-syllable words - in their formula, which was termed "FORCAST." This formula became:

$$\text{FORCAST RGL} = 20 - \frac{\text{number of one-syllable words}}{10}$$

Based on their results, and considering that this variable is probably the most simple count to achieve reliably, the rationale for recommending the one-variable formula seems sound. It should be noted, however, that others have found that the addition of a sentence-length variable adds considerably to predictive power. Kincaid, et al. (1975), for example, found that the addition of a sentence difficulty factor to the word difficulty variable increased the "coefficient of determination" (indicating the degree of shared variance between the predicted grade level and comprehension of the test passage) from 41.6% to 57.2%. This is a sizable increase in power, indicating that the findings of Caylor, et al. might be reexamined to determine if the difference in power outweighs the ease of computation.

Carver (1973c, 1974) also developed a single-variable formula, which he termed the RIDE scale (an acronym for Reading Input Difficulty Estimate scale). The variable used in this formula is simply the average number of letters per work (lpw) in a passage. There are five levels which index RIDE difficulty: Level 1 (up to 4.0 lpw), Level 2 (4.1 to 4.5 lpw), Level 3 (4.6 to 5.0 lpw), Level 4 (5.1 to 5.5 lpw), and Level 5 (5.6 lpw and above). This formula would appear to suffer from the same deficiency as the FORCAST formula. That is, it considers only word difficulty and ignores sentence length. Further, it remains to be seen how this formula would hold up with technical writing containing long, but familiar terms, in view of the fact that it was validated on the Miller-Coleman passages (See Aquino, 1969) using school children. This question should be tested to determine its validity.

Carver based his work on the findings of Bormuth (1966, 1969), which showed that the number of letters per word was one of the highest *single* correlates of the cloze difficulty of 330 100-word test passages. According to Bormuth (1969) only the Dale "Long List' correlated more highly as a single variable than number of letters per word (and that by only .006). It should be noted, however, that these variables both deal only with *word* length/familiarity, and do not touch upon sentence difficulty. It has been shown consistently that word length or familiarity accounts for a greater amount of variance than does sentence length. If one's major goal is to devise a prediction formula containing only one variable, therefore, the rationale for the RIDE scale would be sound. It seems worthwhile, however, to pursue further the potential increase in predictive power with the addition of a sentence difficulty factor

Bormuth (1969) did go beyond the examination of a single variables in isolation. He conducted an extensive correlational study using approximately 170 different variables, of which lpw was one. As a first step, he determined the correlation of each of these variables single with the scaled difficulty level of passages as referred to above. However, he then went on to determine, using multiple regression techniques, which of these variables in *combination* yielded the best predictive power. He arrived at four basic passage length formulas (each of which was computed

against four different criteria). Two of these formulas (termed "unrestricted") employ large numbers of variables and are clearly unsuitable for other than very sophisticated automated computation. (Interestingly, neither of these basic formulas includes lpw as a variable.) To exemplify these formulas, and considering the fact that automation seems obvious to become the rule for the ultimate analysis of difficulty, the "short form" of the unrestricted formula will be described briefly. This formula consists of eight different variables plus the square of two of these, totaling 10 entries in the formula. The eight basic variables are (1) number of Dale Long List words, (2) letters per minimal punctuation unit (and its square), (3) number of "referential repetition anaphora" (and its square), (4) number of numerical nouns, (5) number of derived adjectives, (6) number of common nouns, (7) number of relative clauses, and (8) number of "class inclusive anaphora." It seems obvious that the present state-of-the-art of readability analysis does not permit efficient application of such a complex formula. Its correlation with the mean cloze criterion (.874), however, is exceeded only by that of the even more complex basic unrestricted formula (.889). The question remains (even disregarding computer state-of-the-art) whether these correlations represent diminishing returns over the somewhat lower correlations of the much more simple manual and machine computation formulas described below.

Looking at the letters per word variable alone (as used by Carver) and comparing it to the two Bormuth formulas which include it and other variables, the results are as follows:

(1) Letters per word alone correlated -.721 with mean cloze scores. (For the purpose of this comparison only *mean* cloze score will be examined.)

(2) Letters per word, when combined with letters per minimal punctuation unit in the *manual* computation formula, correlated .808 with mean cloze score (and .833 in cross-validation[3]).

(3) Letters per word, when combined with number of Dale Long List words and words per sentence in Bormuth's *machine* computation formula, correlated .833 with the criterion (and *.920* in cross-validation).

Disregarding the fact that each of the correlations increased in cross-validation, the addition of new variables to the single lpw variable resulted in an increase in observed (or shared) variance of from 14 to 17%. It remains to be seen whether the additional difficulty and possible added unreliability in counting letters per minimal punctuation unit or words on the Dale Long List would offset this rather sizable

[3]Cross-validation was accomplished by applying the formulas to 20 passages of 275-300 words each taken from earlier work by Bormuth (1969).

increase in predictive power. Carver reports correlations of .93 and -.94 between the RIDE scale and Dale-Chall and Flesch formulas, respectively, using the Miller-Coleman passages as criteria. He concludes that, *for these passages,* "...there seems to be little difference between the predictive validity of Dale-Chall, Flesch, and RIDE" (Carver, 1974, p. 12)[4]. These discrepant findings--both of the other formulas consider sentence length *and* word difficulty or familiarity--remain to be explained.

The Fog Count

The Fog Count, developed originally in 1953 by McElroy for the Air Force and described in a *Guide for Air Force Writing* (AFM 11-3), is perhaps the most simple two-variable formula to compute manually. It involves simple counting the number of words of three or more syllables, the total number of words, and the number of sentences and entering them into a simple formula. Rather than discuss the original Fog Count, it would seem more appropriate to address the "recalculated" Fog Count presented by Kincaid, et al.(1975). According to these authors, the revised version is "very similar" to the original except that a different subtracted constant is used to redress a problem of over-estimation encountered with the original. The Kincaid, et al. recalculation was accomplished using 569 Navy enlisted personnel reading material from Navy rate training manuals, at least some of which was technical in nature. In both the original and the revision, long familiar terms are treated specially to avoid the problem of their contributing erroneously to the difficulty of the material. For example, "General Eisenhower" (consisting of a total of seven syllables) is considered as one "easy" word for the purpose of calculating the Fog Count. Other units, such as numbers and some abbreviations are treated in a similar manner. It would seem that a similar procedure could be applied in dealing with technical terminology *known* to be familiar to the user of a given piece of material. Here a problem enters that is related to the first of the basic questions posed above. Before the Fog Count (or any other formula) can be made to take account of such terms, the terms themselves must first be *identified.* This is not so easy as one might first suppose. In the field of electronics, for example, a basic dictionary of terminology for comprehension of written material could be obtained from an examination of books in the area. But to whom will this dictionary apply? Will it be sufficient to have only one such dictionary? If there is only one, will it apply equally as well to the difficulty of material designed for basic electronic training (e.g., "A" school) and the senior petty officer on the job? Considering these questions, the research involved would seem to be a difficult, but not insurmountable effort.

---

[4] Carver points out, however, on the basis of later work, that the Flesch and Dale-Chall formulas seem to be "consistently" better than others.

And once the dictionary(ies) is (are) constructed, there is the question of whether the additional effort required in applying them is accompanied by a significant and practical gain in our ability to gear materials to the desired level of difficulty. With regard to the Fog Count, and presuming that a dictionary of technical terms proves worthwhile, the present advantage of easy manual calculation will probably be reduced. The effective use of a dictionary of technical terms will virtually demand the automation of the predictive device. The Fog Count, in manual form, however, might even then be of value for a "rough" difficulty analysis where automation is not available. Reference to a "short list" of, say the 25 most common terms, could be accomplished during a manual count, and such words treated as "easy" for the purpose of computation. Further reference to this problem will be made with regard to the Dale-Chall formula below and in the discussion of "production" of readable writing.

The Dale-Chall Formula

The Dale-Chall formula (Dale & Chall, 1948) is a two-variable formula which is based on average sentence length and the number of words in the passage *not* on the "Dale List of 3,000 Easy Words." (This "Dale List" was compiled by asking fourth graders which of a number of words they "knew.") To the best of this author's knowledge, the Dale-Chall (D-C) formula has not been validated in a military setting. It is discussed here because it has consistently been found to be highly valid when compared with other formulas in non-military settings, and because it offers a "model" for construction of a technical dictionary or lookup list such as discussed above. Again, presuming that the appropriate technical terms have been identified, such terms could possibly be simply added to the Dale list. Then, when encountered in text (since they are on "the list"), they would be considered "known" or "easy" and not contribute to the difficulty index of the passage. If research indicates that a technical dictionary does improve our prediction ability for military writing, the D-C would be recommended for use due to its "track record" of validity.

Other Methods of Measuring Difficulty

The above discussion has considered only the prediction of readability by basically "statistical" methods. Mention should also be made of two other methods for determining the readability and comprehensibility of written material. The first of these involved a judgment of the quality of material by persons who read a passage and compare it to some standard or to their own experience of job reading requirements. Carver (1974) examined this procedure using his "Rauding Scale." A major advantage of this method over that of formulas is that style, sentence structure, etc. can be assessed by the reader, a task which formulas are unable to carry out. In Carver's approach, judges who qualify on the "Rauding Scale Qualification Test" are asked to judge sample passages and categorize them according to a set of six "anchor passages." Using the average rating of three such judges, a grade level for the target passage is determined. In short, Carver found the correlations between Rauding Scale judgments and

the RIDE, Flesch, and D-C formulas to range from .74 to .84. These moderately high correlations indicate that human judges are taking into account characteristics of writing similar to those used in calculating the formula scores. It seems equally obvious, however, that human judges are considering "something else" in their ratings; what this might be one can only surmise. Carver (1974) hypothesizes that the Rauding Scale reflects "...the difficulty of the ideas or concepts in a passage," and indicates when "...choppy sentences and inappropriately inserted little words" make material harder to read and understand (than a statistical estimate would indicate). The major disadvantage of the use of human judges in the assessment of readability and comprehensibility is simply availability--availability of time and of personnel. For such a procedure to be used effectively with the vast amount of material with which the Navy is concerned would seem to be prohibitive in the extreme.

Another method for determining the quality of written material is the "cloze test" referred to above in the context of readability criteria. There seems to be considerable disagreement as to *what* the cloze test actually measures. Some say it measures readability; others use the terms comprehensibility, comprehension, or understanding. Kincaid, et al. (1975), for example, state that "...subjects were tested for their *understanding* of the selected passages using the cloze procedure" (p. 3, italics added). Sticht (1975c) says "...the cloze test provides a valid measure of reading *comprehension*" (p.20, italics added). A number of writers (including Taylor [1953], who developed the procedure) refer to the cloze technique as a measure of *readability*.

If the operational definition of readability is taken to be the index provided by a readability formula, the cloze test will *not* consistently measure readability. The results of a cloze test on the "scrambled" version of "Mary Had A Little Lamb" (for example) would probably bear little resemblance to the low difficulty predicted by a formula. Clearly, this is an extreme case. When writing of fairly high quality is examined there is consistently a moderate to high correlation between cloze performance and readability as predicted by formula. But one would expect, as the general "quality" of writing decreases, this correlation would correspondingly drop. It does not seem wise to use as a measuring device such a "rubber yardstick."

A correlation is also consistently found between cloze scores and "comprehension" or "understanding." It is fairly well established that a cloze score of about 40-45% corresponds to a comprehension test score (after reading the material) of about 75%. But again, it can be shown that there are limiting factors. This author has conducted a relatively modest pilot study in which cloze tests were administered on a single passage of about 200 words. Subjects were then asked to state as precisely as possible what the passage was all about. Only 3 subjects out of the approximately 50 tested was able to provide even a "reasonable" approximation of the passage content. Yet the mean cloze score over all

subjects was approximately 40%. This would seem to be at least an indication that the subjects were not "comprehending" what they were reading. It is anticipated that a follow-up study will examine the differences in performance when (1) a *title* descriptive of the content of the passage is provided and (2) a traditional test of comprehension is administered following the cloze test. Again, since there are clearly limits beyond which the cloze-comprehension correlation will not hold up, should we not be more precise in speaking about cloze in general?

It is held here that close measures *comprehensibility*. Klare, Sinaiko, and Stolurow (1972) take a similar position when they say, "... cloze measures the relative comprehensibility of intelligibility of written material." To reiterate an earlier statement, readability and comprehensibility are *not* one and the same. Comprehensibility entails such factors as writing style, sentence structure, and expectancy for familiar words. The first two of these factors cannot be detected at all by available readability formulas. The latter can be to some extent, although the identification of those words which *are* familiar to a given audience has not yet been satisfactorily accomplished. Yet these factors (and undoubtedly others, as well) are critical to performance on a cloze test. Multiple "embedding," for example, would most likely cause cloze performance to deteriorate greatly. Consider the following sentence:

> "This is the malt the rat the cat the dog teased killed ate that lay in the house that Jack built."

In this example--a syntactically correct sentence with a relatively low level of difficulty according to formula--a person encountering the sentence for the first time would likely have a very difficult time with a cloze version. Style--that undefinable characteristic of writing--will cause even greater problems. There seems to be little doubt that the smooth flow of ideas, with few departures from the "standard" syntax, enhances the probability that the word which occurs to a reader for a given "set" will be the correct one. In short, it is considered that the cloze test should be considered a *supplement* to a readability formula. Given that a passage is readable (as predicted by a formula), the close performance of a set of subjects similar to the intended audience of the passage should ensure that the material is also comprehensible.

Assessing Job Reading Requirements

Formulas, judgments, and cloze scores all clearly play a role in determining the level of ability required by a user of written materials. They do not, however, give any indication of the reading ability required for a specific *job* or even of the *reading* required for that job. The Navy Personnel Research and Development Center (NPRDC) and the Human Resources Research Organization (HumRRO) recently set out to seek answers to these important questions. This research effort involves the direct assessment of job incumbents' ability to read and understand job reading

materials. It involves a structured interview on the job site in which the men are asked to report what specific reading they have undertaken in the previous day or so that was directly connected with their work. When a man reports that he used a specific portion of a specific publication to perform a specific job, he is asked to retrieve that same information in the presence of interviewer. From this interview data, a Navy Reading Task Test (NRTT) battery will be constructed, consisting of materials which job incumbents reported that they actually read, and calling for the reader to answer questions of the type job incumbents had in mind when they used the materials. The battery will then be scaled on a sample of recruit personnel of known reading ability but not possessing job-specific skills and knowledges. Once the reading ability required for each of the tasks is determined, they will be submitted to job incumbents for judgments of the criticality of each and the frequency with which each is performed. The final product will be an inventory of various types of tasks and the reading level associated with each. The inventory can then be used for establishing general reading requirements for *many* Navy training or job situations. In short, the inventory will make it possible to determine the nature of reading tasks in a rating, the criticality of each, the frequency with which it is performed, and an estimate of the reading ability required in order to perform it.

It can be seen that, in one sense, this is an alternative to the use of readability formulas in assessing the difficulty of materials designed for a specific job. It has the advantage, however, of determining reading abilities required for specific materials known to be *used* on the job, whereas using formulas, some (or perhaps much) of the material sampled is not actually used by job incumbents (such as "theory of operation" for example). Put another way, the fact that written materials is "designed" for a particular job does not necessarily mean that it is required to perform that job. Needless to say, the NRTT method does not supplant the need for formulas, judgments, or cloze tests. When any new material is sent into the field, it is obviously necessary to ensure, as far as possible, that the material will be satisfactory in terms of readability and comprehensibility. Whether it will be *used*, these tests will not say. The ultimate test of its useability must come from the man on the job.

In addition to providing information as to the reading ability required to perform particular jobs, it is hoped that information regarding "readership" will also emerge. That is, does a gap between reading ability and difficulty of required job reading materials influence the extent to which men report *using* those materials. In discussing a similar study for the Army, Sticht (1975c) reports that "...the more able the reader, the greater the reported use of printed materials" (p.52).

Summary

Summarizing the area of prediction of readability and comprehensibility, it seems clear that not just one, but a number of approaches should be used to promote high quality written materials. Readability is a basic

characteristic and, in its operational sense, must be assessed according to a formula. But this is not sufficient. The comprehensibility of the material must be assured by means of cloze tests or human judgments, if possible. Which formula should be used to predict readability, and which technique(s) should be used to assess comprehensibility are still matters open to question.

PRODUCTION OF READABLE WRITING

General

Predicting the difficulty of written material "after the fact" is a major problem, but equally or more important is the problem of producing "readable writing" in the first place. For this latter task there is no "formula" which one can directly apply. Writing is an ART as well as a SKILL. While it is probably impossible to tell someone how to "do art" there should be some techniques by which the writer's *skill* can be enhanced. This is the concern of this section, and in general, can be conceptualized as a three-part problem: (1) Can a writing style guide be developed which will serve as a "job performance aid" for good writing? (2) Given that a writer is both skilled and artistic in his work, can Military Specifications be provided him so that his product conforms to our needs? and (3) How can we best provide the writer with "tools" to ensure that specified standards are being met?

Readability formulas play a part in the production process, but as pointed out so clearly by Gunning (1968), "...formulas are tools, not rules...warning systems, not formulas for writing." Production is not simply the opposite side of the coin from prediction, although the two processes are inextricably related. The two processes can be thought of as iterative; one writes, then assesses the difficulty of the writing, then rewrites, then reassesses, etc. The interaction between the two will be examined more thoroughly below in the contexts of specific writing aids.

Research in the area of readable writing does give some cause for optimism, but for caution as well. There have been a fairly large number of studies which have shown clearly that experimental manipulation of writing variables can make a significant difference in comprehensibility. In general, these experimenters (many of them conducting PhD dissertation research) have taken original passages and written easier and/or harder versions as indexed by readability formula scores. In examining these sources, the problem is one of specifying precisely *what* was changed to make the readability different. Most have reported that they modified word and sentence difficulty and little else, presumably because these are the two variables common to most readability formulas. Most writers would argue, however, that such changes are either not sufficient or are too simple-minded for such a complex task. It is here that a word of caution is in order. There seems to be little doubt that mechanical shortening

of words and sentences will do little, if anything, other than improve readability formula scores, with scant chance of improving the reader's comprehension. At the extreme, it is clear that a writer with ulterior motives could certainly produce writing that gets a *better* readability score and yet be *less* comprehensible than the original version. It seems clear, then, that in the research studies mentioned above the modification of surface variables (i.e., word length, sentence length, etc.) have incidentally resulted also in changes to deeper (or casual) variables. This is a research question which is virtually untouched and yet which is critical to the understanding of how to make writing more comprehensible.

Style Guides

The first facet of the problem of readable writing as mentioned above--style guides to aid the writer--would profit immensely from the research just suggested. The guidelines put forth by existing style manuals are based largely on intuition, with little, if any, empirical base. It may be that such manuals are of no help to writers. They have existed in profusion in our libraries for years, and yet we are still suffering from inferior written materials. One must conclude from this that such guides are either not used by writers of military materials or that the information they contain is inadequate to the task of telling one "how to write." Information provided this author by Dr. G. R. Klare (personal communication) indicates that the problem lies, at least in part, with the latter of these possibilities. Klare reports that in reviewing 15 source books (10 written specifically for technical writers) the agreement among authors as to specific suggestions was quite low. The suggestion "use short words" (for example), which one might expect all authors to agree upon, was mentioned in only *two* of the 15 books sampled. Outright disagreement was found for such alternatives as "be concise" versus "be complete," and "keep paragraphs short" versus "vary paragraph length." In view of this apparent uncertainty among "experts" on writing, it is not surprising that much of our material is unacceptably difficult. It is felt that progress is being made in this area with the publication of style guides specifically for the military writer. Under contract with the Naval Sea Systems Command, for example, BioTechnology, Incorporated produced a guide titled "Requirements and Criteria for Improving Reading Comprehension of Technical Manuals" (Post & Price, 1974). This guide contains 17 "tests" for improving the quality of technical writing in the three areas of Organization, Technical Communication, and Readability. The seven tests dealing specifically with readability and comprehensibility, in abbreviated form, are as follows:[5]

---

[5] This manual assumes that the writer is aiming at the ninth-grade level; this should be considered in interpreting the tests. An answer of "yes" to any question indicates adequate readability (at ninth grade) on that test.

1. Heading Review. Do approximately one-half of the subparagraphs have headings? Is material within paragraphs consistent with its heading?

2. Topic Sentence Check. Is the heading clear? Does the heading cover about three or four topic sentences or key points?

3. Words Per Paragraph Count. Do paragraphs average no more than 45 to 60 words? Are key points highlighted if the paragraph must be longer?

4. Words Per Sentence Count. Do the sentences average 20 words or fewer? Have compound sentences and complex sentences been avoided?

5. Syllables Per Word Count. Does the material average about 1 1/2 syllables per word? Have short words been used whenever possible?[6]

6. Equipment Nomenclature Count. Is any unfamiliar nomenclature either defined in the text or called out on an accompanying pictorial?

7. Layout Review. Has double-column format been used? Is each graphic contiguous with the text in which it is discussed or referenced?

It is believed that this manual (which also contains specific guidelines for correcting deficiencies, a large number of "rewrite practices," and generous illustrations of "good" and "bad" material) can be of valuable assistance to the technical writer. However, it is vital that it be experimentally tested to verify this assumption. Work is currently underway which involved the use of this guide in rewriting a technical manual and comparing the rewritten version with the original in terms of its effectiveness.

Military Specifications

A second major area of concern in the production of readable writing is that of communicating the needs of the user to the technical writer. This is a question of the clarity, conciseness, and completeness of Military Specifications (Mil-Specs) and Military Standards (Mil-Stds). With regard to readability and comprehensibility, it is the opinion of this writer that present Mil-Specs and Mil-Stds offer little useful guidance for the writer. Other matters aside, it is clearly apparent that no test of readability or comprehensibility has been applied to *these* documents themselves. One sentence picked more or less at random from one basic Mil-Spec contained 47 words of which 21 words were composed

[6] The authors suggest that since manuals deal with technical terms which cannot be eliminated, these terms should not be included in the count of syllables. This procedure should, however, be experimentally tested.

of three or more syllables. Other similar examples abound. But more specifically, with regard to specification of readability and comprehensibility standards, most or all such documents are woefully inadequate. One, for example, states that "As a general guide, the level of writing should be for a high school graduate having specialized training as a technician in military training courses."[7] Even those of us in the field of readability/comprehensibility (much less the engineer/writer) don't really know what a typical high school graduate "looks like." And taken literally, manuals written to this prescription (if it were possible) would automatically be excluded from use in at least some training courses, because the trainee would not yet have the "specialized training" required to understand the manual. It would be comparable to the employment advertisements which specify "only experienced need apply," when the "experience" can be gained only through the employment.

Another, more basic Mil-Spec gives somewhat more explicit directions. It states: "Narrative text (those pages that consist of not less than 200 words in consecutive sentences per page) shall conform to the following readability standards: The average sentence length (ASL) shall not exceed 20 words...The average word length shall not exceed 1.60 syllables... The percent personal sentences (PPS) shall not be less than 15% of the total."[8] The particular figures specified are reasonable; if such standards were achieved, the readability formula score would be at approximately the 9th or 10th grade level. But, one might wonder if some writers read any further than where it says "...200 words in consecutive sentences..." Certain persons responsible for producing technical manuals have been heard to say that "by definition" their writing never *has* more than 200 words per page in consecutive sentences. This not to imply that producers or writers of manuals intentionally write poorly. Quite the contrary, one suspects that they take pride in their work just as any professional does. But the fact remains that much of our technical writing *is* too difficult for many of the intended users. More specific guidance, and perhaps more rigid quality control, would appear to be necessary conditions for rectifying this problem. And accompanying such specifications, it is felt that we must provide the writer with tools to meet our standards--the third general problem in the production of readable writing.

---

[7] MIL-M-24100B, Manuals, Technical: Functionally Oriented Maintenance Manuals (FOMM) for Equipment and Systems, Jan 74.

[8] MIL-M-63000C (TM) Manuals, Technical: General Requirements for Manuscripts, Dec 1960.

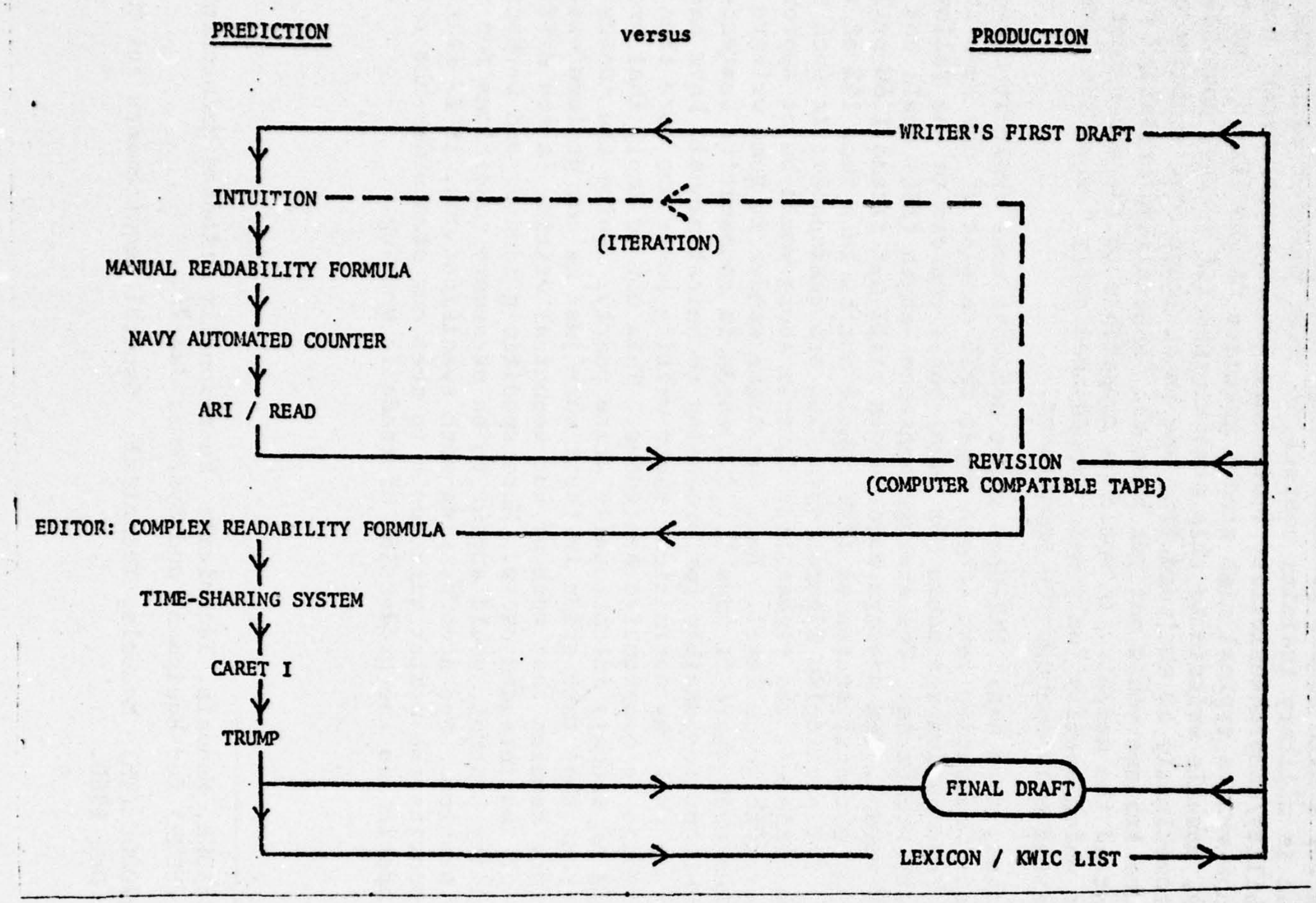


Figure 3. Prediction of Difficulty versus Production of Readable Writing

Tools For Readable Writing

What tools could be provide for writers? In general, the answer involves the various methods of assessing readability and comprehensibility discussed in the previous section. As indicated earlier, prediction and production go hand in hand. Figure 3 illustrates the iterative process that is involved. Included in this figure are some examples of the tools that might be provided the writer in order to produce an acceptable product. Basically, we see the writer producing his "first draft," relying at the outset primarily on intuition in predicting its difficulty. At some point, however, a more precise assessment of the writing should take place. This might take the form of a readability formula (which the writer will be trained to use) computed manually. As pointed out earlier, the formulas which are currently available, from RIDE and FORCAST to the probably more precise Flesch (Kincaid) and Dale-Chall, are all relatively simple to calculate by hand. As our understanding of the problems specific to military writing (e.g., long, but familiar technical terms) become better understood, and as cost-effective and more efficient methods of analyzing readability and comprehensibility are developed, it would seem appropriate to automate the process of prediction to the greatest possible extent. Several possibilities present themselves in this regard, ordered in successively more sophisticated techniques.

One such technique is the Navy Automated Counter (NAC) developed for use in the CNETS analysis of Navy Rate Training Manuals (RTMs) (Biersner, 1975; Bunde, 1975). To reiterate briefly, this device is composed of a stylus, which, when pressed to a working surface, trips a micro-switch, which in turn activates a counter. Used for counts of syllables, words, and sentences for the Flesch and Kincaid formulas, it would be equally as useful for computing virtually any other formula. This could be a cost-effective aid which writers would probably find more acceptable than using simply paper, pencil, and fingers. This "acceptability" would probably give greater assurance that accurate counts are in fact being made in the field. Bunde (1975), in a test of this device, found that it reduced the time to compute the Flesch formula by 46% over manual counts, with equivalent reliability. From the point of view of this writer, such a device might be considered as an *interim* technique. It is recommended, however, that more sophisticated procedures be planned for future use.

At a somewhat higher level of sophistication are those readability analysis techniques which utilize a specially modified electric typewriter. These are the Automated Readability Index (ARI) and the Reading Ease Assessment Device (READ). Given material to be analyzed during any typing stage of the production process, these devices (which are similar in design) allow the writer to determine the difficulty of his writing as it is typed. Essentially, these devices take account of the number of "strokes" of the machine to count the number of letters per word and the number of words per sentence to be entered into a modified Flesch RE formula. They both require typing skill, however, which may be a drawback

for some writers. In addition, their cost may be prohibitive (although not exorbitant) for situating at the many locales at which manuals are produced. If however, the cost is warranted, and if skilled personnel are available for their use, they represent a step advanced from the manual counting device discussed above. But in addition to the above factors, these devices are limited in that they provide *only* readability scores. As will be seen below, improvement of technical writing overall actually demands much more.

At a higher level of potential, let us consider two possible systems--one marketed by Scientific Time-Sharing, Incorporated, and the other developed at Harvard University.[9]

The first of these systems incorporates a modified RE computation, in which a sample of material can be input to a computer terminal which is connected via telephone line to a central computer. Material is typed at a remote terminal and, on signal, an almost instantaneous analysis of the writing is printed out. A seemingly advantageous characteristic of this technique is its ability to reduce the impact of long technical terms on the analysis. When inputting the material, if a familiar but long technical term is encountered, it is counted as any other word at its first appearance--both as a discrete word and as a polysyllabic word. When subsequently encountered, it can be enclosed in brackets which excludes that occurrence from the overall analysis. The effectiveness of this technique, however, depends entirely on the ability of the typist to recognize those words which are of such interest and bracket them in an efficient manner. Unlike the ARI or READ devices, however, it is considered that the time-sharing system offers the potential for such features as storing of technical terms in a lexicon and future *automatic* exclusion of these terms from the analysis. It is considered that future R&D effort might well include the investigation of such potential.

The second of these computerized techniques is referred to as a Computer Aided Revising, Editing, and Translating system (CARET I), and was developed by Klare, Rowe, St. John and Stolurow (1969). Like the time-sharing technique discussed above, CARET I also provides readability analyses after the material to be assessed has been input. It is felt that CARET I has at least two distinct immediate advantages over the former system. First, it provides not one, but *five* separate analyses, including the Flesch, the Fog Index, and the Farr-Jenkins-Paterson revision of the Flesch formula. Secondly, it provides a triple-spaced

[9] No bias in favor of these specific developers is intended. Other similar (and perhaps better) systems may well exist and not have come to the attention of this author.

AD-A034 730 HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA F/G 5/10
READING AND READABILITY RESEARCH IN THE ARMED SERVICES.(U)
SEP 76 T G STICHT, D W ZAPF N00014-76-C-0312
UNCLASSIFIED HUMRRO-FR-WD(CA)-76-4 NL

3 OF 4
AD-A
034 730

printout of the inputted material, indicating for each word the number of syllables, and for each sentence the number of words. The editor/writer can therefore adjust (as necessary) word length and/or sentence length, entering alternatives directly on the printout. And, like the time-sharing system, it is interfaced with a large computer which would permit the compilation of a lexicon and many other possibilities to assist not only in prediction of difficulty, but in writing and rewriting. Figure 3 shows how such a lexicon (as well as a Key-Word-in-Context file) might "feed back" into the various stages of writing. To exemplify the potential of such a system, the entire text of (for example) an electronics manual could be analyzed for the frequency of occurrence of key words at the same time that the material is input for readability analyses. Then not only those words in the lay language, but also technical terms, could be cataloged according to their frequency and this information fed back to writers for consideration either in rewriting or in original writing of new material in the same subject matter area. Optimally, it would appear that the best features of both the preceding systems could be combined to form a procedure that is efficient, effective, and cost-beneficial.

An even more sophisticated system would be one such as the Navy's TRUMP (Technical Review and Update of Manuals and Publications) system. This system is currently undergoing test and evaluation at the Naval Air Rework Facility, Jacksonville, Florida. TRUMP is technologically more advanced than the others mentioned, basically because of its Optical Character Recognition (OCR) and automatic photocomposition capabilities. This makes it possible to electronically scan material to be input, resulting in a reported throughput rate of potentially hundreds of times greater than with keypunching. While no readability analysis for TRUMP has as yet been implemented, it would seem to have the potential for virtually any such analysis as well as all other capabilities mentioned in connection with previous systems. Of particular interest for Navy manuals, TRUMP can automatically process illustrations and complex tables as well as running text. Upon completion of the ongoing T&E of TRUMP, it should be considered as a central feature of any future R&D effort in the area of readability and comprehensibility.

SUMMARY

The underlying theme of this paper has been a plea for reformulation of the basic rationale underlying research in readability and comprehensibility. It recommends that future research be directed at determining which of the many variables involved in the readability and comprehensibility area stand in *causal* relationships with the ability of persons to comprehend the written word. And finally it suggests that present readability formulas, with perhaps some modification, are acceptable when only *readability* is being considered, but that a means must be found to ensure that materials are also satisfactorily *comprehensible*. It seems that the time has come to orient our efforts toward *implementation* of what is already known in the field, and to provide for the military writer the means to achieve an acceptable product considered in terms of the needs of the user.

COMMENTS ON THE PAPER BY CURRAN

George Klare
Ohio University

In commenting on Curran's paper, Klare first reiterated points made by Curran regarding three types of writing for which readability formulas are not useful. First, a readability formula is not useful for predicting comprehension of prose that does not follow the rules of the language; e.g., words strung randomly together or phrases such as "a rose is a rose is a rose". Secondly, a readability formula may not be a good predictor for some prose that *does* follow the rules of the language, but may include complicated embeddings or the like; e.g., "the house that Jack built". Thirdly, readability formulas may not well predict the comprehension of certain types of semi-prose; e.g., highly technical writing, mathematical writing, or poetry. In this case, however, special formulas have been or might be developed. In addition to these three problem areas for readability formulas, Klare indicated that matters or organization, "smoothness of prose style", and related factors limit the effectiveness of a readability formula in predicting comprehension. And, finally, readability formulas do not take into consideration characteristics of readers, such as their background, reading competence, and motivation. In short, Klare agreed with Curran that readability as measured by the various formulas is a necessary, but not sufficient condition, for comprehensibility.

Readability formulas are predictive tools, and one needs to know the conditions under which they predict well and where they do not predict comprehensibility accurately. Readability formulas use predictive index variables, not measures of causal relationships. This fact has several implications. (1) It may be difficult to improve the predictive validity of readability formulas by such things as incorporating a technical term factor in the formula. In one such case, with the Dale-Chall formula, adding technical terms did not alter the predictive power of the formula. (2) We need to continue looking for better criteria of comprehensibility and useability to be predicted by readability formulas. The index variables predictive of comprehension may change depending on the measure of comprehension being predicted. (3) One must be cautious in trying to use what is known about the readability score to produce or to monitor writing. In some cases it is possible to "fool" the readability formula by changing the material to get a higher readability score, while leaving comprehensibility unchanged. Regarding implications of readability research for writing, Klare suggested that if one is going to rewrite material for readability, he or she should work on the material for the least motivated first. Secondly, he or she should rewrite the least interesting of the material before other portions of the material.

RESEARCH ON THE COMPREHENSIBILITY OF AIR FORCE TECHNICAL MATERIALS

Arthur I. Siegel
Applied Psychological Services, Inc.

The Air Force, like any modern organization, depends to a considerable degree on information which is transmitted through natural language encoded in written form. Often the transmission process is ineffective or only partially effective. Increasing the effectiveness of this written word information transfer process is the thrust of most current Air Force programs in the topic area of readability research. Specifically, the Air Force is concerned with increasing the comprehensibility of written materials, developing methods for diagnosing where and why written materials are not comprehensible, and developing automated techniques for assessing the comprehensibility of written materials. Since technical training and education represent a major aspect of the Air Force's peacetime activity, and since a substantial portion of the training process involves transmitting information in written form, it is not surprising that much of this research has been performed in a technical training context. However, there is little reason to believe that the results and implications of prior and current Air Force work are not mostly generalizable to other contexts; e.g., operations and maintenance manuals, bulletin board notices, and technical orders.

Partly because of the technical training orientation and, more importantly, because of definitional problems, much current Air Force work has veered away from media and format problems. There are, however, some exceptions to this generalization (e.g., Davis, 1965). The definitional problem seems to arise from confusion among some researchers between what is meant by "utility" as opposed to the meaning of "comprehensibility." Job aids are judged by the criterion of utility; text is judged by the criterion of comprehensibility. A text may be quite comprehensible, but be of little use in certain operational situations. A job aid may be quite useful in a given operational context and be low on comprehensibility. Similarly, such items as page format, font, and type size may have much to do with reading rate, but have little involvement with comprehensibility. When it comes to encoding language through the written process, media and format problems are mostly related to utility rather than comprehensibility. Comprehensibility and utility should not be confused. The differentiation is not trivial and the Air Force has tried to avoid this confusion.

Subsequent sections of this paper attempt to recapitulate current Air Force trends relative to increasing the effectiveness with which students organize the knowledge encoded in the language appearing in their instructional materials.

Auditory Supplementation

A set of recent Air Force studies has investigated the gain, if any, that can be obtained from supplementing written textual training materials with auditory presentations (Sellman, 1970; Siegel, Lautman, & Burkett, 1974). In this technique, while the reader reads the textual materials, the material is also presented to him through a headset. Accordingly, the approach has been called auditory supplementation. The advantages of the technique appear to be that it might (1) serve to refocus the reader's attention in the case of distraction, (2) help the person who can comprehend spoken language but who possesses a reading disability, and (3) serve as a redundant and noninterfering information source. Present indications from this work are that auditory supplementation may be helpful in increasing comprehensibility in certain situations. Sellman, working at the Air Force Human Resources Laboratory, found that individuals using simplified materials plus tapes scored higher than persons using simplified materials alone. Siegel, Lautman, and Burkett, who also worked under Air Force Human Resources Laboratory sponsorship, partially confirmed Sellman's findings. Although many interactive effects seem to exist, by way of generalization, it seems that auditory supplementation will be helpful when one or several of the following conditions exist: (1) the mental ability of the reader is low, (2) the reading grade level of the reader is low, (3) the materials to be used are of the discursive rather than the "look up" type, and (4) reading materials represent the sole source of the information to be learned.

Comprehensibility Measurement

The Air Force's emphasis in the comprehensibility measurement area focuses on developing new comprehensibility measurement concepts, validating these concepts, and implementing them. Accordingly, the approach consists of a basic research vector, an advanced development vector, and an application vector. Each draws from and is based on the other avenues and considerable interplay occurs. This overall approach seems to introduce considerable economy and cross-fertilization into the cycle.

For a comprehensive review of earlier work relative to comprehensibility measurement, the reader should see the Air Force Human Resources Laboratory report authored by Williams, Siegel, and Burkett (1974). The conceptual emphasis of the present work is based on the conjecture that prior comprehensibility measurement techniques (e.g., Flesch, 1951; Dale & Chall, 1948; and Lorge, 1944) deal principally with what one might call mechanical factors such as quantities of words, sentences, syllables and their occurrences—not with meanings per se. They have been used for some time not only because they measure reading difficulty (specifically, reading grade level) in some sense, but also because they are suitable to relatively easy calculation by hand. However, such measures are not diagnostic. That is, knowing that a given text is at such-and-such reading grade level will tell the author little about how to improve the writing. Moreover, the concepts included in such mechanical measures are atheoretical. They

possess little foundation in language processing or cognitive theory. Additionally, one can use small words and sentences to produce difficult material or, alternatively (given a technical situation), one can use big words and sentences and produce highly comprehensible material. Finally, these measures have been largely validated in high school/grade school situations which have little relationship to adult technical training situations.

The orientation of the Air Force's present work relative to comprehensibility measurement is on the assessment of the cognitive and the language processing load which a text places on the reader. Alternatively, the measures under development seek to assess the amount of thinking a text forces a reader to perform. These measures, accordingly, represent an attempt to reflect and quantify what the text forces to happen inside a reader's head, rather than to determine comprehensibility on the basis of sizes and frequencies of words, length of sentences, and the like. Some support for the contention that cognitive processing aspects represent a salient aspect of comprehensibility comes from the work of Coke (undated), who concluded as the result of a readability study in another context that "...reading rate is sensitive to the amount and/or type of cognitive processing required by a reading task." Additional support seems to derive from the ongoing work of Davis (1975) at the Air Force Technical Institute, relative to the effects of mode of expression on comprehensibility.

Within the intellective load structure, two separate, but related approaches are under investigation: (1) an "intellective load" analytic approach, and (2) a "language processing" approach.

The intellective load analytic approach is based on and drawn from the structure-of-intellect model of intellective function developed by Guilford and his associates (Guilford, 1967; Guilford & Hoepfner, 1971). This model is the result of over 20 years of work which empirically isolated and defined 120 intellective functions. A logical choice was made among the 120 Guilford functions on the basis of rational relationship with textual comprehensibility. The Guilford functions selected, and on which textual comprehensibility assessment techniques were developed, are cognition of semantic units, cognition of semantic relations, evaluation of symbolic implications, and divergent production of semantic units. Full descriptions of these metrics are found in Siegel and Wolf (1975) and in Siegel and Bergman (1974). The Siegel and Bergman report also presents the results of an initial validation of these measures. Cross validation and predictive studies are currently being performed.

The language processing approach is based on concepts drawn from current psycholinguistics. Metrics have been developed to reflect each of the following: Yngve depth, morpheme volume, transformational complexity, center embedding, left branching, right branching, and complement deletion. These metrics are fully defined, described, and elaborated in Siegel and Wolf (1975) as well as in Lambert and Siegel (1974). The Lambert and Siegel report also describes the methods, procedures, and

results of an initial validation study relative to these psycholinguistically oriented measures. Currently, cross validation and predictive studies are also being performed relative to metrics.

Automated Measurement of Comprehensibility

The Air Force's goal in the automated comprehensibility measurement sphere is the development of an on-line technique which will handle the logic and calculational sequences required to determine the selected structure-of-intellect and psycholinguistically oriented measures described immediately above. The computer program which will achieve this goal is currently being specified and set in logic form. It is anticipated that formal programming will start during early months of 1976.

Obviously, a part of the computer processing of text for publication is the "typing" or optical reading of the text. Some text editing programs also use a computer-stored dictionary. With this capability, a computer can accomplish functions such as automatic hyphenization, page numbering, indexing, page layout, spelling checks, centering of headings, and the like. The comprehensibility measurement program which is being developed could become the back end of the more routine text handling procedures now available or could stand on its own.

Some concept of the organization of the comprehensibility measurement computer program can be gained from Figure 1. Input and interaction with the program can take place in the batch mode or through a terminal. The operating subroutines represent those program modules necessary for setting up and performing a textual analysis. The semantic subroutines are those subroutines necessary for performing the calculations underlying the selected structure-of-intellect and psycholinguistically oriented metrics. A parsing subroutine is included because of the dependence of certain of these metrics on sentence structure and syntactic attributes. Results will be presented, at user option, by sentence, block of sentences, and total text. Block and sample size are also variable under user control. Output will be in usual printed format, on cards, on tape, or on the terminal (in the interactive mode).

The utility subroutines include a dictionary update routine (needed because many technical words may not appear in the usual computer-stored dictionary, a teaching subroutine (to teach the user how to use the total program), and a number of statistical subroutines (to summarize data).

The support subroutines represent those usual procedures needed for an effective program (checkout, alarm, etc.), a subroutine to plot directional trends, and norms for comparing a text under analysis with a comparison set of texts. The required norms have already been developed.

Figure 2 further details the analytic sequence in its current form.

The reader who is interested in an earlier Air Force approach to the automation of comprehensibility analysis will find the work of Smith and

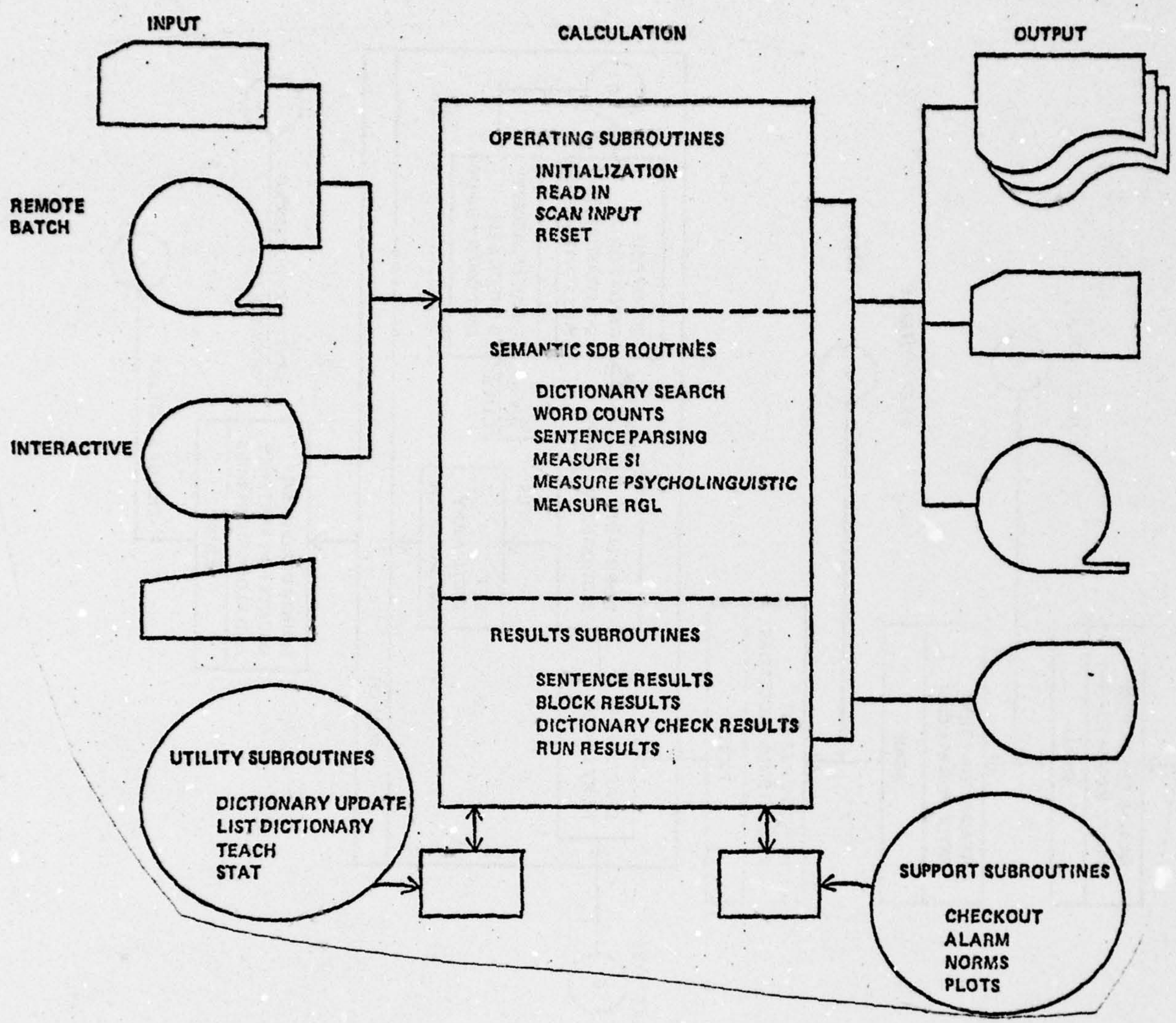


Figure 1. Overview of the Comprehensibility Measures Program

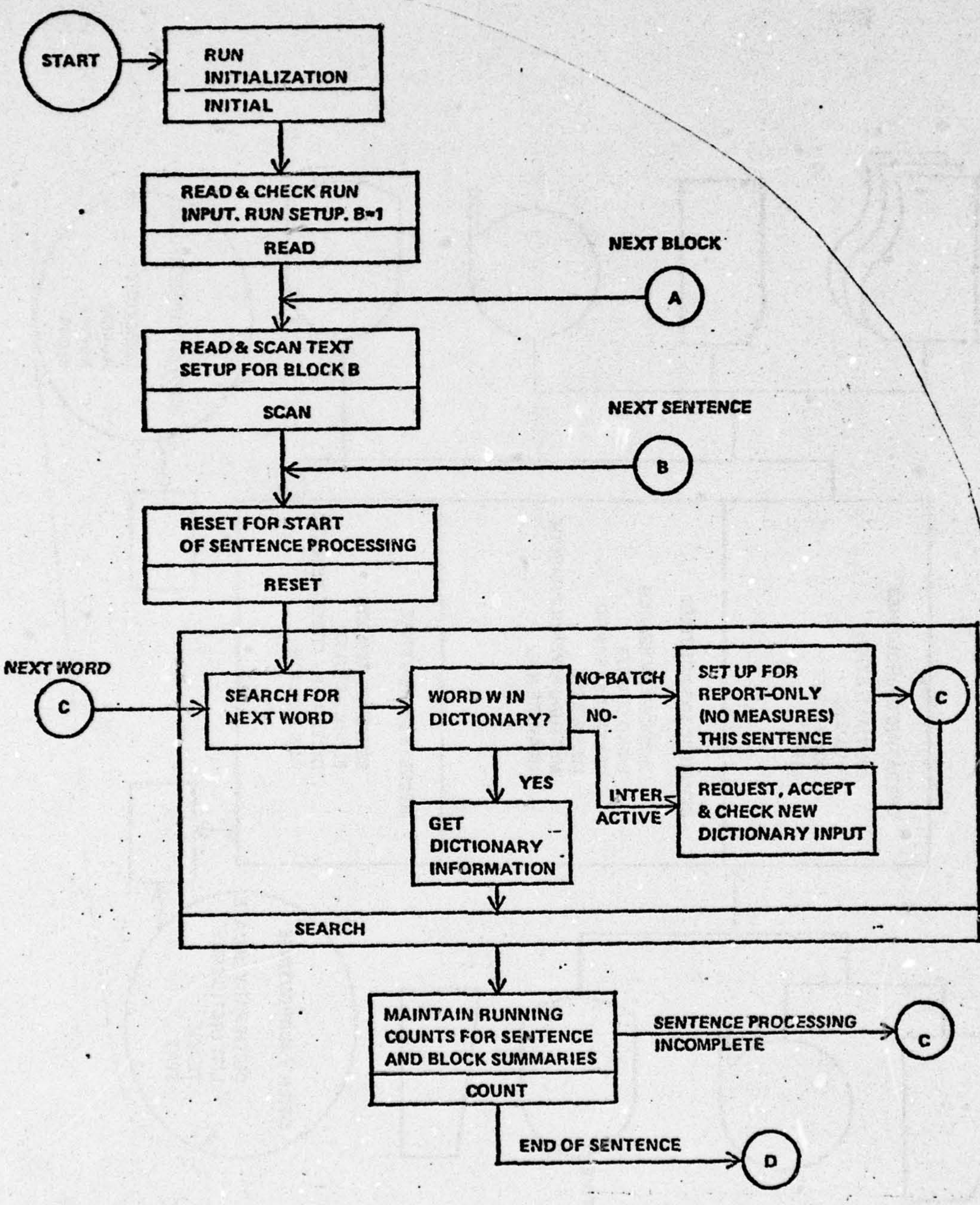


Figure 2. Analytic Sequence

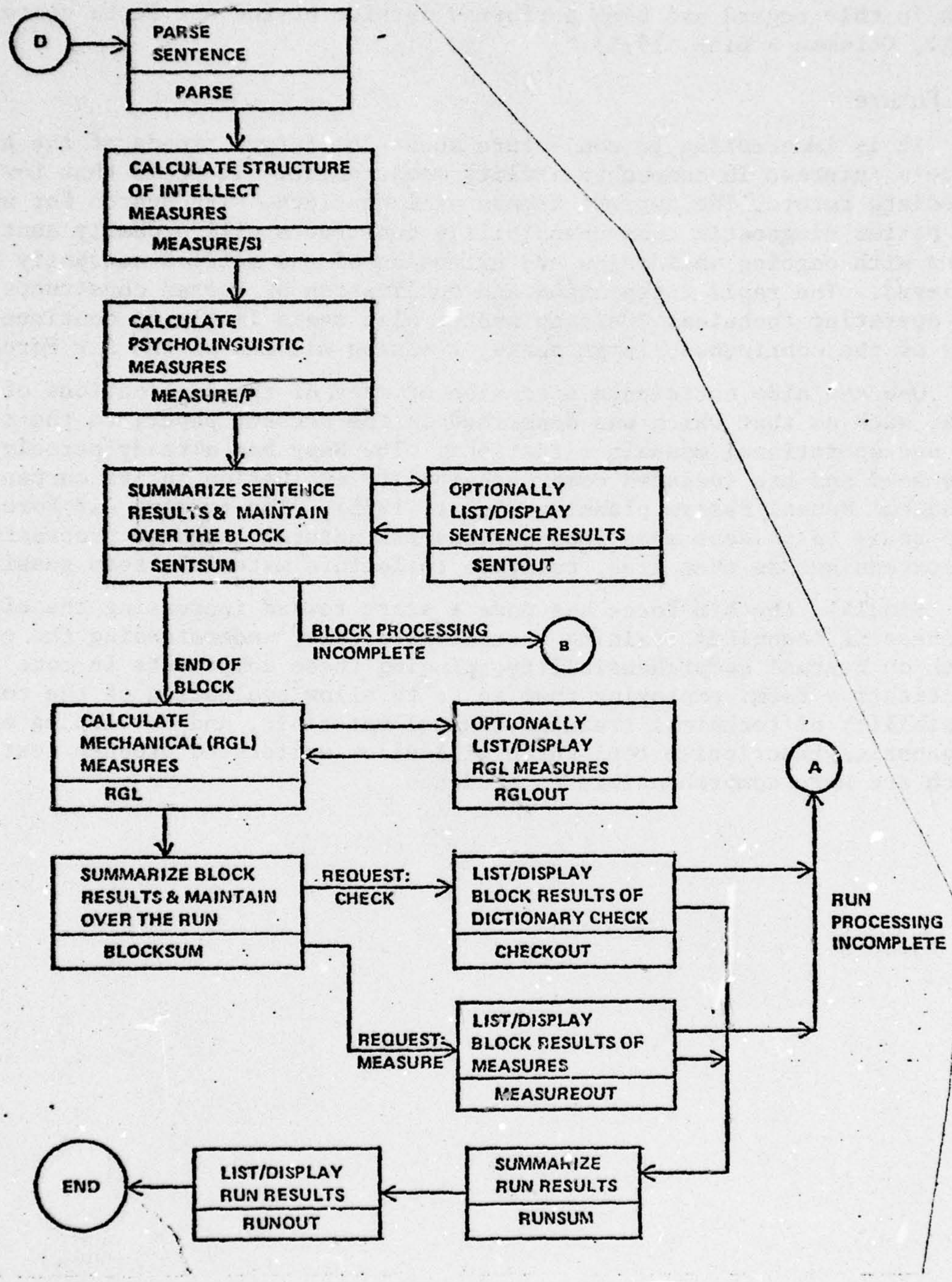


Figure 2. Analytic Sequence (Continued)

Kincaid (1970) and that of Smith and Senter (1970) of relevance. Other work in this regard has been performed outside of the Air Force context (e.g., Coleman & Liau, 1975).

The Future

It is interesting to conjecture about the future trends of the Air Force's interest in comprehensibility measurement. It seems that for the immediate future, the current trends will continue. The search for new and better diagnostic comprehensibility constructs will probably continue, along with ongoing validation and extension of the metrics currently being achieved. The rapid integration and application of tested constructs into the operating technical training system also seems likely to continue in view of the continuous, large scale, training mission of the Air Force.

One can also anticipate extension of many of the implications of work, such as that which was described in the present paper, to the technical and operational manuals situations. The Navy has already perceived this need and has included comprehensibility evaluation in its current Navy Technical Manual System planning (Sulit, 1975). The current Air Force work also seems to possess some implications for natural language processing, and extensions in that area, relative to lecture material, seem possible.

Finally, the Air Force has made a start toward increasing the effectiveness of technical training textual materials, understanding the components of textual comprehensibility, placing these components in some quantitative form, employing them so as to allow evaluation of the comprehensibility of technical training textual materials, and developing a diagnostic/prescriptive tool which will allow writers to produce texts which are more comprehensible to students.

COMMENTS ON THE PAPER BY SIEGEL

Ernst Rothkopf

Bell Telephone Laboratories

As a preliminary opening comment, Rothkopf expressed his opinion that the military appears to have a concern for the human use of written materials, and to have shouldered the responsibility for making such documents more effective. Rothkopf felt that civilian publishers, universities, and school systems have been relatively irresponsible in this effort, or perhaps they have recognized that the need existed, but saw no means for implementing any action in that direction. The military work may provide guidance for the civilian sectors on the practical ways in which people produce written materials for useful purposes.

Regarding military work on the comprehensibility of materials, Rothkopf made some general remarks regarding the relevance of the military work to those who are engaged in more fundamental research on human learning and on language processes. The point made was that the work in job aids and other documents that are supposed to support job behavior reveals the folly of accepting the word "comprehension" as a primitive term, the way it has been accepted as a primitive term by many psycholinguists and students of cognition. When simply studying, learning, and recalling information from text, it is possible to never know that one has not understood. But when the printed material is supposed to take one to the next stage in a procedure or sequence, one quickly recognizes that language is not all that informative, and that understanding is a very complicated thing.

Directing his remarks to Siegel's paper, Rothkopf made two major points. First, he agreed with Siegel that internal measures of text (such as word length and sentence length) are interesting and ought to be explored. However, he stressed the idea that the variables counted in readability formulas, such as word length in the Flesch formula, have a certain construct validity as measures of more psychologically interesting variables, such as familiarity, which in turn have an effect on learning, retention, and so forth. Thus, word length variables are not without psychological meaning. The second major point concerned the comprehension measures which Siegel reported were planned to be used. Rothkopf suggested that a parsing program which took semantic considerations into account but which operated only on measures internal to the document would not provide an adequate evaluation of the document's comprehensibility. He submitted as an area for research the idea that it is meaningless to think about evaluating a document by itself; the only way that a document can be evaluated is with reference to some outside criteria. The outside criteria have to come from the purpose of the document or some other pragmatic

concern. Thus, a document characterized as being comprehensible on the basis of some general measure derived from internal measurement may be a completely inadequate document for its intended purpose, because it can contain wrong or irrelevant information. Rothkopf cited a study in which 25 or 30 structural variables were examined, and not a single one had the predictive power as the amount of irrelevant material that was in the text. But the irrelevant material cannot be detected without reference to some purpose external to the document. For this reason, Rothkopf saw the need to somehow bring the content variable into the evaluation of documents.

U.S. ARMY RESEARCH AND DEVELOPMENT ON READABILITY AND USEABILITY OF PRINTED MATERIALS

Richard P. Kern

U.S. Army Research Insititue for the Behavioral and Social Sciences

INTRODUCTION

The U.S. Army's literature research and development efforts will be described in this paper in terms of work addressing three problem areas: (1) the reading difficulty level of the literature, (2) the content and the orientation of the literature, and (3) assessment of the utility or useability of the literature when placed in the hands of the intended user.

READING DIFFICULTY LEVEL OF ARMY JOB AND TRAINING LITERATURE

The first major research which examined reading difficulty levels of Army job and training literature was stimulated by the Department of Defense announcement of Project 100,000 in 1966. The implementation of Project 100,000 was expected to produce an influx into the Army of large numbers of men with marginal or low literacy skills. The research initiated by the Army in 1968, under contract with HumRRO, had the broad objectives of developing information required to identify literacy skill requirements of common, high density, Army jobs. Based on this information, the most feasible approaches could then be developed to enable the Army to effectively train and utilize the lower aptitude Project 100,000 personnel. The resulting program of research carried out for the Army under HumRRO Work Units REALISTIC, READNEED, and FLIT is summarized in Caylor's paper in this volume. Because of the early influence and continued relevance of these studies to Army R&D on job and training literature, certain key features of the readability and readership findings will be summarized here.

Reading Skill Level of Personnel and Reading Difficulty Level of Manuals

It was expected that one obvious point of difficulty for the marginally literate personnel would be the Army literature they were expected to use in learning and performing their job. Research (Sticht, et al., 1971), comparing reading skill levels of low aptitude job incumbents with the reading difficulty level of their job literature confirmed this expectation. However, this research also indicated that many of the high aptitude job incumbents could be expected to have difficulty in using the same literature.

Figure 1, reproduced from the Work Unit REALISTIC report (Sticht, et al., 1971), indicates that reading and learning from Army literature is a difficult task not only for low aptitude personnel but, also, for many high aptitude personnel. This figure presents a visual comparison

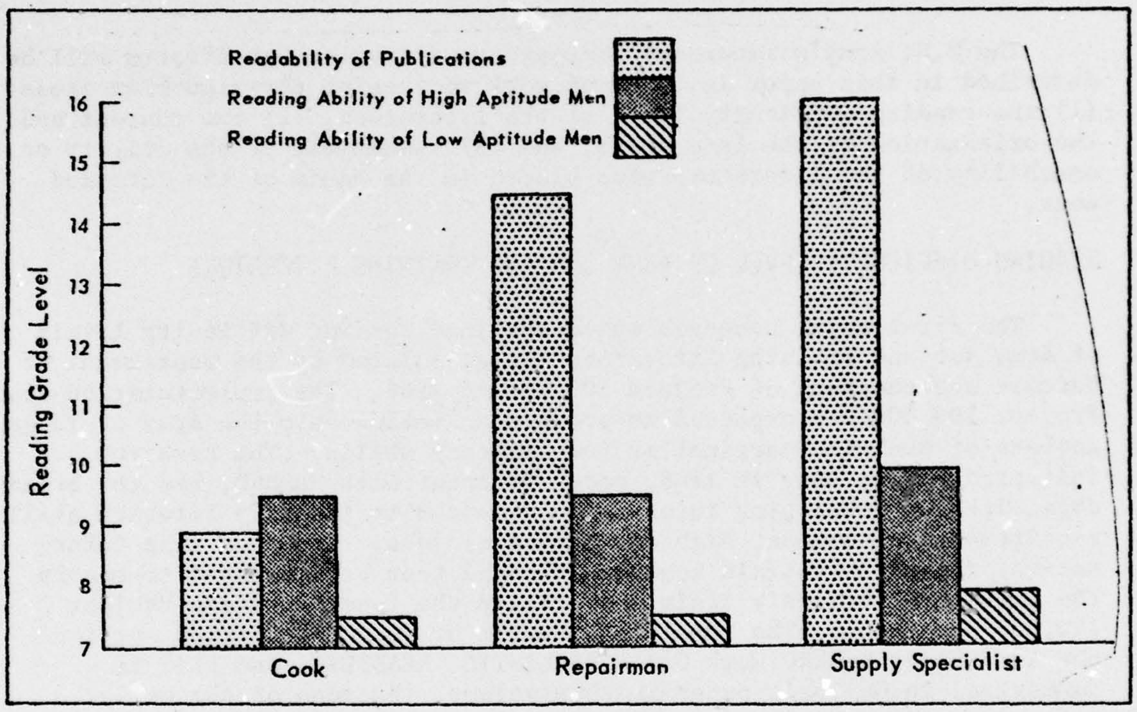


Figure 1. Readability of Publications and Reading Ability of Job Incumbents.

between the average reading difficulty level (modified Flesch formula) of manuals cited by incumbents and supervisors as being used on the job, and the average reading skill levels of incumbents in three high density Army jobs. Note that for repairmen and supply specialists the average difficulty level of the manuals exceeds the average reading skill levels of even the high aptitude incumbents by 5 to 6 grade levels. The reading materials used by the cooks, however, appear to be reasonably geared to the reading skill levels of the high aptitude cooks and, while difficult, possibly still within the range of the low aptitude cooks. The lower difficulty levels of the cooks materials reflect the fact that the bulk of the materials cooks reported using on the job were their job aids, the recipe cards.

A second, related set of results from Work Unit READNEED are presented in Figure 2 (Sticht, et al., 1973). In this case the manuals studied were identified from the list of study materials prescribed for the soldier's use in preparing for his annual Primary MOS Enlisted Evaluation Test in each of the seven MOSs. The FORCAST formula, developed in Work Unit READNEED (Caylor, et al., 1973) for the Army population and Army technical materials, was used in estimating reading difficulty level of these manuals. Average reading skill level is indicated by the vertical lines for three groups of Army personnel: Army Preparatory Training (APT) graduates for FYs 1968, 1969, and 1970; Category IV job incumbents; and non-Category IV job incumbents tested during this and a related research program (Vineberg, Sticht, Taylor & Caylor, 1971). The APT program, at the time this research was conducted, contained remedial reading instruction in the context of a six-week general educational development program. Use of the study materials to prepare for their MOS Evaluation Test would clearly be a formidable if not impossible task for the APT graduates and the Category IV personnel regardless of which of these seven MOS they might be assigned. It is equally important to note, however, that these study references are written at such a high level of difficulty that even the higher aptitude (non-Category IV) personnel will be expected to have problems in reading and learning from them.

Reading Skill Level of Personnel and Use of Manuals on the Job

Studies conducted in the fields of journalism and advertising would lead us to expect that as the reading difficulty level of material increases, readership or use of the material would decrease. From this we might extrapolate and assume that job incumbents with, for example, 10th-grade reading skill levels, will be more likely to use job manuals written at, say, a 12th-grade reading difficulty level than will job incumbents with lower reading skill levels.

Estimates of the extent to which incumbents of differing reading skill levels used manuals or other printed material in carrying out their job duties are reproduced from a Work Unit REALISTIC report in Figure 3 (Sticht, et al., 1973). These estimates were obtained during interviews conducted with job incumbents. Using a version of Flanagan's Critical Incident technique (Flanagan, 1954), each incumbent was asked

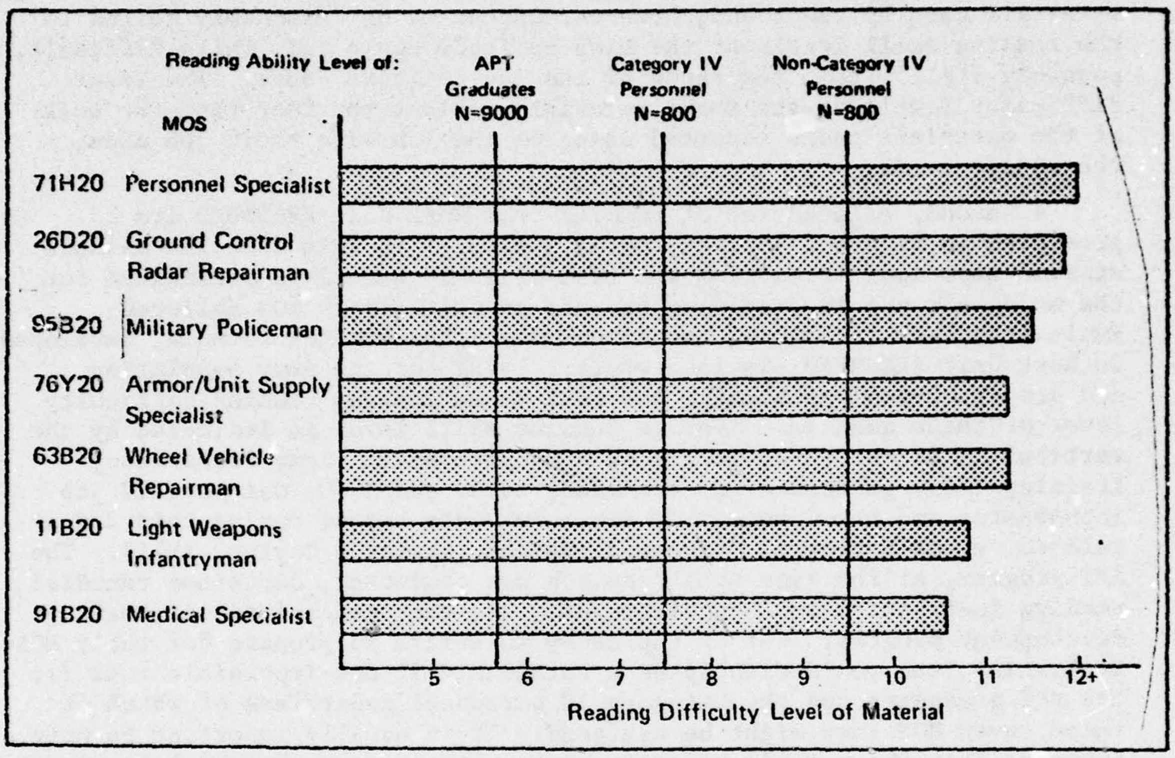


Figure 2. Average Reading Difficulty Level of Materials in Seven MOSs.

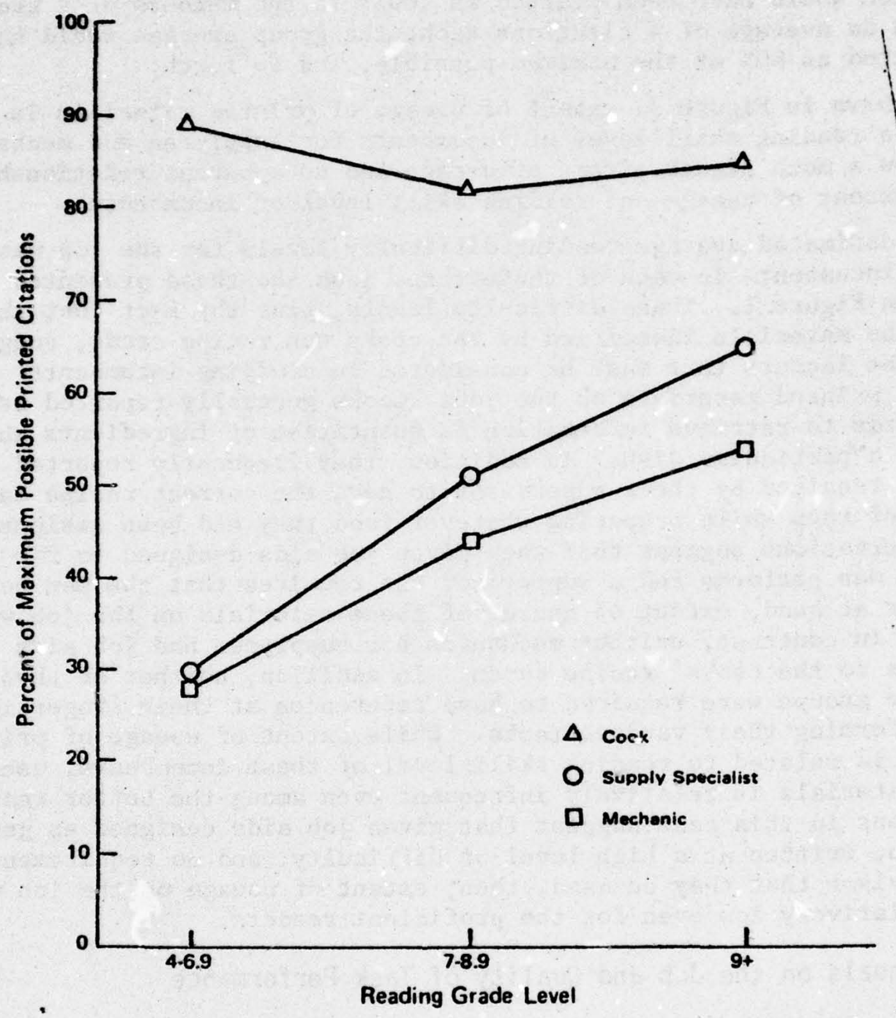


Figure 3. Reading Ability and Use of Job Reading Materials.

to give five instances during the past month when he had consulted printed materials in connection with carrying out his job. Extent of useage is plotted in Figure 3 as percent of maximum possible number of citations a group could give. For example, if each man in a group had given the maximum of five instances when he had consulted printed materials, this information would have been plotted as 100%; if the members of a group had given an average of 4 citations each, the group average would have been plotted as 80% of the maximum possible, and so forth.

As shown in Figure 3, extent of useage of printed materials is related to reading skill level of incumbents for supplymen and mechanics. Cooks show a much higher extent of useage and no apparent relationship between extent of useage and reading skill level of incumbents.

The estimated average reading difficulty levels for the job materials cited by incumbents in each of these three jobs are those presented earlier in Figure 1. These difficulty levels, plus the fact that the bulk of the materials identified by the cooks was recipe cards, suggest some of the factors that must be considered in studying incumbents' useage of printed materials on the job. Cooks generally reported using recipe cards to retrieve information on quantities of ingredients while preparing a particular dish. In addition, they frequently reported that they were required by their supervisor to have the correct recipe card in front of them while preparing whatever food they had been assigned. These observations suggest that when given job aids designed to fit the tasks the man performs and a supervisor who requires that the man have these aids at hand, extent of useage of these materials on the job will be high. In contrast, neither mechanics nor supplymen had job aids comparable to the cooks' recipe cards. In addition, neither of these latter two groups were required to have references at their fingertips while performing their various tasks. While extent of useage of printed materials is related to reading skill level of these incumbents, use of printed materials is relatively infrequent even among the better readers. Observations in this case suggest that given job aids designed as general references, written at a high level of difficulty, and no requirement by the supervisor that they be used, then, extent of useage on the job will remain relatively low even for the proficient readers.

Use of Manuals on the Job and Quality of Task Performance

The research reviewed in the preceding section indicated that among repairmen and supplymen, use of manuals on the job was positively related to reading skill level of the incumbents. However, is there any evidence that use of manuals improves the quality of task performance? Data related to this question were obtained in Work Unit REALISTIC. Sticht, et al. (1971) describes this research as follows:

> In that research, job incumbents were administered three- to five-hour Job Sample tests in which Repairmen actually repaired vehicles and Supply Clerks worked in simulated offices filling out forms and counting equipment. In these Job Sample tests,

job manuals were available so that men who were being tested could use the manual if they wanted. Figure 4 shows relationships between reading ability, use or non-use of manuals, and performance on the Job Sample tests for Supply Specialists and Repairmen. It is apparent that men who use manuals did better than those who did not use manuals regardless of reading ability. Furthermore, of those who used manuals, performance was better for the more highly literate men.

Initial Approaches to the Readability Problem by Army Commands

In view of the research findings on reading difficulty levels of Army training literature and average reading skill levels of enlisted personnel, Headquarters, United States Continental Army Command (CONARC) established the 9th-grade reading difficulty level as a maximum difficulty level goal to be used in preparing and reviewing literature intended for entry level enlisted personnel. In addition, CONARC directed that a review be made of the reading difficulty level of all manuals produced by CONARC schools. The 20 CONARC service schools, using the FORCAST readability formula, calculated average reading difficulty levels on each of 470 Field Manuals (FMs) and Technical Manuals (TMs) for which they had responsibility. These analyses revealed that approximately 94% of the CONARC manuals exceeded the 9th-grade reading difficulty level and, in fact, approximately 65% had an average difficulty level of 11th-grade or higher. A frequency distribution showing the number of manuals at each reading grade difficulty level is shown in Figure 5 (CONARC letter dated 29 Jan 1973).

In reviewing these results CONARC recognized that not all of their 470 manuals were either designed for or required by enlisted personnel at the entry skill levels. However, examination of the readability of their literature highlighted the fact that literature management policies in effect at that time did not require specific identification of job, rank, and skill level characteristics of the population for whom a given manual was primarily or exclusively intended. Thus, readability analyses focused attention on the management of literature production and the need for designing literature to meet the needs and abilities of specifically identified groups of users. Research initiated by the Army Research Institute (ARI) to assist CONARC in providing guidance and assistance to their literature management and production agencies is described in the next section of this paper.

The Army Materiel Command (AMC) initiated studies to develop readability standards for use in guiding design and preparation of the many hardware operation and maintenance manuals they are responsible for procuring. This work (Army Materiel Command, 1970), initiated before the development of the FORCAST formula, resulted in the following recommended readability standards:

1. Average sentence length in narrative text should not exceed 17 words.

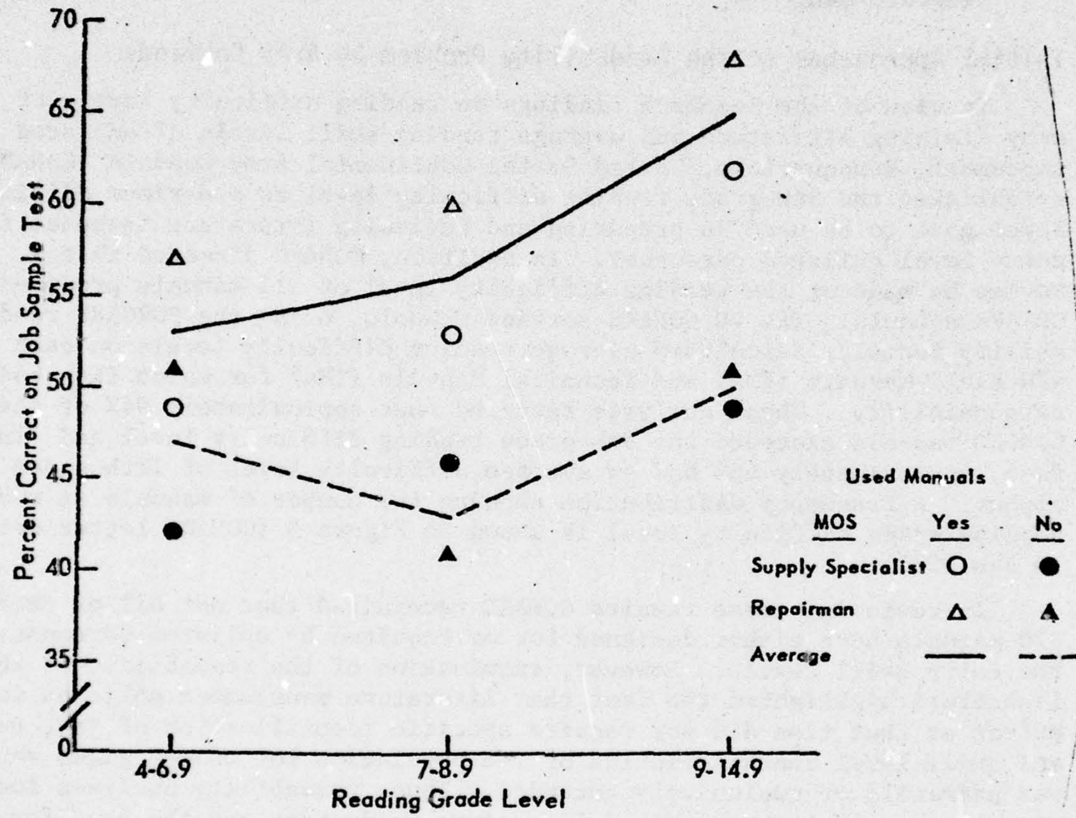


Figure 4. Job Sample Test Performance as a Function of Reading Ability and Use of Technical Manuals.

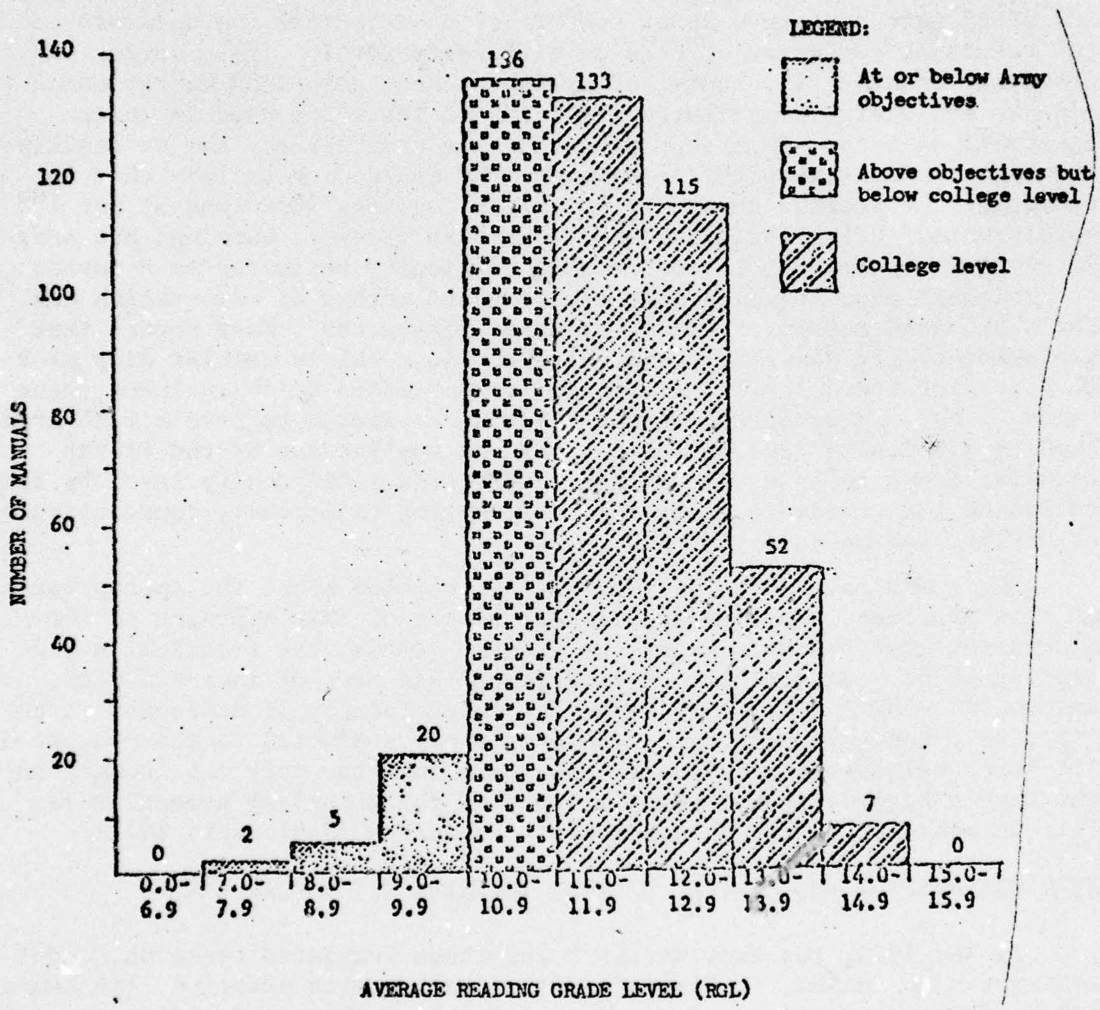


Figure 5. Distribution of Manuals by Average Reading Grade Level.

2. Average word length should not exceed 1.5 syllables per word, excluding technical nomenclature (mandatory) words.

3. At least 20% of the sentences in the narrative text should be personal sentences.

In establishing reasonable standards for equipment manual writers, AMC staff were concerned about the effect of technical nomenclature on the resulting estimates of reading difficulty level. These words generally consist of a number of syllables and, according to estimates made by AMC staff, constituted about 1/3 of the words used in their equipment manuals. On the premise that the user already has or quickly acquires familiarity with these words, AMC decided to exclude them in computing the average number of syllables (average word length) per 100 word sample. Using three different formulas (Flesch, Gunning, and Army-Dale) they reported computing reading difficulty estimates on a number of equipment manuals using both the standard method of computation and their proposed method of excluding manadatory words. They report that exclusion of the mandatory words resulted in a fairly regular drop of 3 to 4 reading grade levels as compared to estimates which included these words. Thus, an equipment manual which is estimated to have a 12th-grade reading difficulty level based on standard application of the Flesch formula, drops to an estimated 8th-grade reading difficulty level by simply excluding the mandatory words and substituting additional, consecutively occurring, non-mandatory words.

There are, of course, differences of opinion about the appropriateness of this approach. Perhaps the weakest aspect of this approach is the assumption that the user knows, or quickly learns, the technical nomenclature words. Studies of job-experienced and new, or inexperienced, incumbents' familiarity with technical nomenclature, if conducted in the Army, have not been reported. One such study conducted in the context of Air Force helicopter maintenance, suggests that the only job incumbents who have a high degree of familiarity with the technical nomenclature used in manuals are the manufacturer's engineers (Jablonski, 1971).

USER ORIENTATION AND CONTENT OF ARMY TRAINING LITERATURE

In May 1973, the Army Research Institute initiated research, under contract with HumRRO, to identify user problems with existing literature and user requirements for revised training literature. In addition, research was also initiated under this same contract to identify and develop guidance to assist Army writers in preparing literature appropriate to the user's job needs and reading skill level.

Survey of User Requirements for Army Training Literature

The survey of user requirements for training literature was conducted within active Army units by means of a structured interview (Showel and Brennan, 1976). Personnel surveyed included students, instructors, and training managers assigned to Army Training Centers and Service Schools.

In addition, interviews were conducted with first line supervisors responsible for conducting instruction at the Company or Battery level in Infantry Divisions, and those responsible for training management at the Company, Battalion, and Brigade levels.

Findings applicable to the general body of training literature include complaints that information contained in FMs and TMs is frequently out of date, scattered through several different manuals, and presented at such a high level of generality that it is difficult to determine how it applies to specific situations. In addition, students reported that the literature was difficult to read, dull, too wordy, and that it was difficult to locate needed information in the manuals.

In general then, these interviews indicated that reading difficulty level of the literature, while recognized as a problem, was not considered as great a problem as the highly generalized orientation of the content and the failure to organize and bring together content relevant to the user's needs into one compact, easily accessible source.

Development of Guidance for Army Writers

In addition to the survey of users of Army training literature, researchers conducted interviews with Army writers and editors. The purpose of these interviews was to find out what type of training or experience writers brought to their job, what type of guidance and support they received, how they perceived the purpose or the function of the literature they prepared, and what they thought a writer's guidebook should include to be of help to them (Kern & Sticht, 1974).

Information obtained during these interviews indicated that writers were generally selected on the basis of their previous or current performance as classroom instructors. In many cases the writing assignment was added duty for an instructor.

In general, the purpose or function of the training literature, from the writer's perspective, was to provide official, detailed, reference support for the highly generalized lecture-training approach that was still prevalent in the schools at that time. Writers were provided with little or no instruction regarding users, purpose, or scope when tasked to write a new manual. Guidance generally was simply to write a manual on topic "X". Consistent with this guidance, writers viewed their job as that of assembling, organizing, and recording all that was known about a *topic*. They assumed that if a manual contained a reasonably complete exposition of the topic, it could serve as a general text or reference source and be used by anyone interested in the topic regardless of the specific nature of the reader's job duties or information needs.

Most of the writers interviewed said they did not use writer's guides of any type. Guidance regarding format and organization was obtained by using previous editions of the same manual or a related manual as a model and preparing the new literature to look like this model. Whenever possible they cut and pasted from the earlier manuals as well as from current editions of related manuals. Editorial review of the writer's

draft was generally limited to grammar, proof reading, and conformance to format policies. Readability formulas or other methods for testing the difficulty level of the writing were not being used.

The above observations indicated that if writers were to produce literature to support the Army's emerging demands for performance-oriented training, they needed help in re-orienting their perception of the purpose and functions of training literature. They would also need guidance in how to research, develop, and prepare this literature so that it was appropriate to the job-related information needs and reading abilities of their targeted users. It was recognized that a guidebook could not be expected to accomplish these objectives unaided. It was thought, however, that it could provide a basis for further training efforts by literature management.

The *Guidebook for the Development of Army Training Literature* (Kern, Sticht, Welty & Hauke, 1976) was designed to capitalize on the writers' reported tendency to use existing literature as a model for writing style, organization, and format rather than devoting his time to the study of technical writer's guides. It seeks to acquaint writers with the user orientation of performance-oriented writing by presenting contrasting examples of topic-oriented and performance-oriented writing. It also presents detailed guidance on how to identify a primary user for a proposed manual and how to proceed in developing, organizing, and writing the material to meet this users' information needs.

The Guidebook deals with the difficulty level of writing in the following ways:

1. Presenting short reading tests which the writer can use to compare his reading level with the average reading levels of different groups of enlisted job incumbents.

2. Presenting many examples of difficult writing ("Before" examples) obtained from a wide variety of Army manuals. These "Before" examples are accompanied by a contrasting version ("After" example) rewritten to reduce reading difficulty level and to shift the orientation of the material from a topic-orientation to a performance-orientation.

3. Providing directions on the use of the FORCAST formula to check drafts for possible rewrite needs.

In addition to the guidance described above, illustrated guidance is provided on identifying when and how illustrations can be helpful and on design of tables of content and indexes to make it easier for users to locate information in a manual.

This guidebook, using an unconventional approach for a writer's guide, has been well received by Army writers. It is currently undergoing test-use and review by the U.S. Army Training and Doctrine Command (TRADOC).

Development of User-Oriented Literature by the Army

Recognizing the need for providing literature or other media that would more effectively support the individual soldier's efforts to learn his job and perform his job, several Army agencies initiated development of user-oriented literature. The U.S. Army Combat Arms Training Board (USACATB) served as a catalyst and coordinator of these efforts for TRADOC and also performed the editorial and graphic design work for many of the early publications. These early publications, aimed at the individual combat soldier, were small booklets focusing on knowledge and skills the soldier would need to perform a relatively specific combat duty; for example, how to destroy enemy tanks or how to use camouflage, cover, and concealment. Text was kept brief and colored cartoon figures were depicted as narrators. Efforts to catch the individual soldier's attention included, in addition to the use of cartoons and color, titles such as "To Catch a Tank: Big Game Hunting Made Easy". Each publication contained a post-paid reply card requesting the soldier to provide feedback regarding his reaction to the publication: whether or not he liked this type of publication; whether or not he found the information useful; and, whether or not he would like to see other subjects presented in this manner. Informal reports by USACATB personnel indicate that these early publications were received with enthusiasm by the targeted enlisted audience. The use of cartoons, however, came under criticism by some officer personnel who considered them undignified and demeaning.

Since these initial efforts to produce user-oriented literature, a number of new equipment operator manuals have been produced through the combined efforts of the training (TRADOC) and the equipment (AMC) Commands. These manuals were designed to limit content to the operator's needs; to present content by means of illustrations closely integrated with concise, supporting text; to avoid referencing from one page to another; and to emphasize warnings by use of a second color (Braddock, 1975).

Crew members' duties and crew drill procedures for crew-served weapons have traditionally been published in an FM while a separate TM described the weapon, its operation, and maintenance. AMC and TRADOC have recently completed and are preparing to field test a single operator's manual containing all of the information crewmen need to operate and maintain a given crew-served weapon (Braddock, 1975).

The above are only a few examples of ongoing Army efforts to revise traditional topic-oriented literature and prepare it as performance-oriented literature developed to meet the users' information needs while in training as well as on the job.

ASSESSING USEABILITY OF ARMY LITERATURE

As described at the beginning of this paper, early research focused the Army's attention on the reading difficulty level of its manuals. Subsequent research and studies by Army management broadened this initial focus to include the problems of content and the absence of a

user-orientation in the design and presentation of the information.[1] Now that various efforts are being made to develop literature intended to surmount these problems, the Army's attention is being further extended to focus on the problems of verifying the useability of new manuals once produced and placed in the hands of the user.

The Army's long standing practice of sending new manuals to appropriate schools and field commands for review and comments cannot provide the type of information needed to verify useability in the hands of the intended user (Highlander, 1975, Kern & Sticht, 1974). In looking for an effective approach to this problem, Smith (Smith, H. L., 1975) has proposed adopting the conceptual approach used for developing and verifying the Army's Training Extension Course (TEC) materials.

TEC represents the Army's new approach for developing multi-media, performance-oriented, self-paced instruction which can be exported to Army field units. These courses have utilized audio/visual, audio only, and printed (programmed text) media. The audio/visual version of TEC is considered the "workhorse" of the program and may be used either on an individual basis or a small group basis. The developmental approach used in TEC (Smith, H. L., 1975) consists of first identifying critical tasks to be trained; secondly, analyzing each task to determine optimum procedures and to identify performance standards; thirdly, using the performance standards as a basis, developing performance tests; and fourthly, developing training to teach the performance test. A major feature of the developmental cycle is the requirement that draft lesson materials be subjected to user testing to determine their effectiveness in communicating to the intended soldier population. Soldiers' responses to self-study questions contained in the lessons, their performance on pre- and post-tests, and their responses obtained during interviews are studied to identify the need for revision of the materials. If revised, materials are again subjected to user testing. This cycle continues until the materials meet previously specified criteria for acceptance.

Smith (1975) reports that applying this same approach to development of Technical Manuals will result in manuals which present otherwise complex information in the form of step-by-step instructions. Furthermore, by using this approach to concurrently develop training and Technical Manuals, it is believed that fewer tasks will be selected for training

---

[1] It should be noted that researchers dealing with the design of job and training literature to support equipment maintenance personnel had been focusing on the problems of identifying content requirements and developing user-oriented literature for a number of years prior to research reported in this paper. Fink (1967) presents an excellent summary of conclusions, relevant to the issues, based on research conducted for the Army up to 1967.

since they can be adequately covered by step-by-step instructions presented in the manual. The Army Materiel Command (AMC) and Training and Doctrine Command (TRADOC) have recently joined forces to conduct a demonstration project using this approach for concurrent development of improved technical manuals and training programs for the tank turret mechanic. Smith describes the major steps in this demonstration project as follows:

> <u>Step I Job Analysis</u> - This step is taken to insure a full and accurate definition of the job, as it is actually performed in the unit.
>
> <u>Step II Task Analysis</u> - This step will identify the specific activities to be performed, such as the detection, removal, or replacement of a faulty part on the tank turret. In this analysis optimum procedures will be developed by skilled technicians. It is through these optimum procedures that the need for time consuming decisions will be eliminated from the mechanic's job.
>
> <u>Step III Structure Documentation and Training</u> - At this point (Step III) a strategy for training and documentation will be developed. Given the tasks and optimum procedures developed in Step II, Tasks are identified for training site and media selection. This step will identify tasks to be included in the training program, or in the manual.
>
> <u>Step IV Develop Documentation and Training</u> - The methods for conveying the maintenance information are developed. The actual lesson materials and technical manuals will be fleshed out.
>
> In this last step, perhaps the most important of all, will be validation. It is here that the training and technical data are soldier tested to insure ability to perform the job.
>
> . . . .
>
> The demonstration will serve as a learning vehicle for the Army. A full evaluation will be made to document the benefits and lessons learned so that application to other systems can be made. (Smith, H. L., 1975, pp. 104-105.)

Smith does not elaborate on how the new manuals are to be user-tested; that is, whether, as in TEC, the soldier testing will be part of the developmental cycle, or, as in the current Air Force studies (Klesch, 1975), the soldier-testing represents a summative evaluation of the relative cost and performance effectiveness obtained under different manual and skill level conditions.

Past Army Research on Manual Effectiveness

Several past Army research and development efforts have produced new manuals as a component of experimental training programs (e.g., Gebhard, 1970; MacGaslin, Woodruff & Baker, 1959). In some instances the manuals were developed to serve both as a training aid during training and as a performance aid once on the job. Criteria used in

these evaluations have been performance on selected job tasks while using the same job aid or manual (experimental or conventional) the soldier used during job training (experimental or conventional). However, since these manuals were developed as a component of the overall training package, the evaluations conducted have focused on the effectiveness of the overall training package and not on the manuals per se.

The one reported Army research effort, of which this writer is aware, that compared effectiveness of a new manual against an existing, conventional manual also used performance on several selected job tasks as the criteria (Rogers & Thorne, 1965). In this study, new graduates of a missile maintenance course were carefully paired and effectiveness of an experimental troubleshooting manual was studied by comparing the performance of those using the manual with the performance of those using the conventional manual and notes used during their training.

Performance testing of the type referred to above is expensive. It is considered necessary to R&D projects that are seeking to establish the effectiveness of new techniques or approaches. However, it is too cumbersome and expensive to incorporate into an operational literature quality control program. In addition, evaluation studies using performance tests as a criterion usually cannot be conducted at the individual's work site and generally require that he be brought to a special testing area and perform under various other special conditions. As a result, in this type of study we learn very little about the operational utility (or useability) which the manual possesses for the man on the job.

Suggestions for Future Research on Useability of Printed Materials

We need to develop relatively inexpensive ways of assessing the useability of printed materials and of identifying design features that make significant differences in useability for given user groups. But, before we can make progress in this regard we need to identify what we mean by the term "useability" and the conditions under which it can be assessed. For example, perhaps the term "useability" can best be considered as a global term used to subsume the following characteristics of the printed material when evaluated in relation to a given user located in his expected work or training site:

1. Design of manual for convenient work-site storage.

As Highlander (1975) points out, size of a manual should be chosen based on its purpose and how it will be used. While this seems obvious, it is a consideration that is apparently frequently overlooked. Manual designers and developers are not at the work site, According to Highlander, manuals for individual weapons which provide no reasonable storage space should probably be pocket-sized. However, a manual for a combat vehicle, which is required to carry a logbook, should take advantage of the larger logbook size since it can be stored with the logbook and always be with the vehicle.

2. User's knowledge of, or training in, how to use the manual.

People generally learn how to read the different type of information displays they commonly encounter. Initial performance with a new type of information display, e.g., the first encounter with a logic tree, will likely produce confusion and, if useage is on a voluntary basis, may result in avoidance of the manual. Studies of various new types of job aids have either incorporated them into the training program, or, as in the case of the Air Force (Klesch, 1975), provided short periods of training on how to use the job aid. Army researchers have also recognized this need (Shriver & Trexler, 1966).

3. Ease of locating specific procedures or information in the manual.

This is another user problem area cited by Highlander (1975). While part of the problem may be lack of knowledge and practice in using indexes, Figures 6-A and 6-B illustrate how this can be made a tedious task or a relatively simple task.

4. Ease of reading or comprehending specific procedures or passages of information.

The problem of reading difficulty level have been discussed earlier in this paper. With the shift to performance-oriented printed materials, however, narrative text has been sharply reduced in favor of what are, frequently, annotated illustrations. Figures 7-A and 7-B illustrate this contrast in design of the information display. What effect does this change have on the user's comprehension of the information presented? We can predict comprehension of the printed narrative in Figure 7-A by using readability formulas. How can we assess and predict comprehension of the information presented in Figure 7-B?

5. Ease of transforming information or procedures to the form required for task performance.

Improving ease of transforming information or procedures to the form required for task performance is the core objective of performance-oriented literature. Detailed analyses of job tasks are carried out to provide the basis for preparing this literature. Can the extent of achievement of this objective be assessed from the perspective of the user?

6. Completeness of task-relevant information.

This characteristic splits out in two ways. One way concerns inadequate information given in a procedure. For example, compare the information presented in Figure 7-A with the corresponding content in Figure 7-B. If you are a totally inexperienced tank gunner but capable reader, which would you find most helpful? The second way concerns

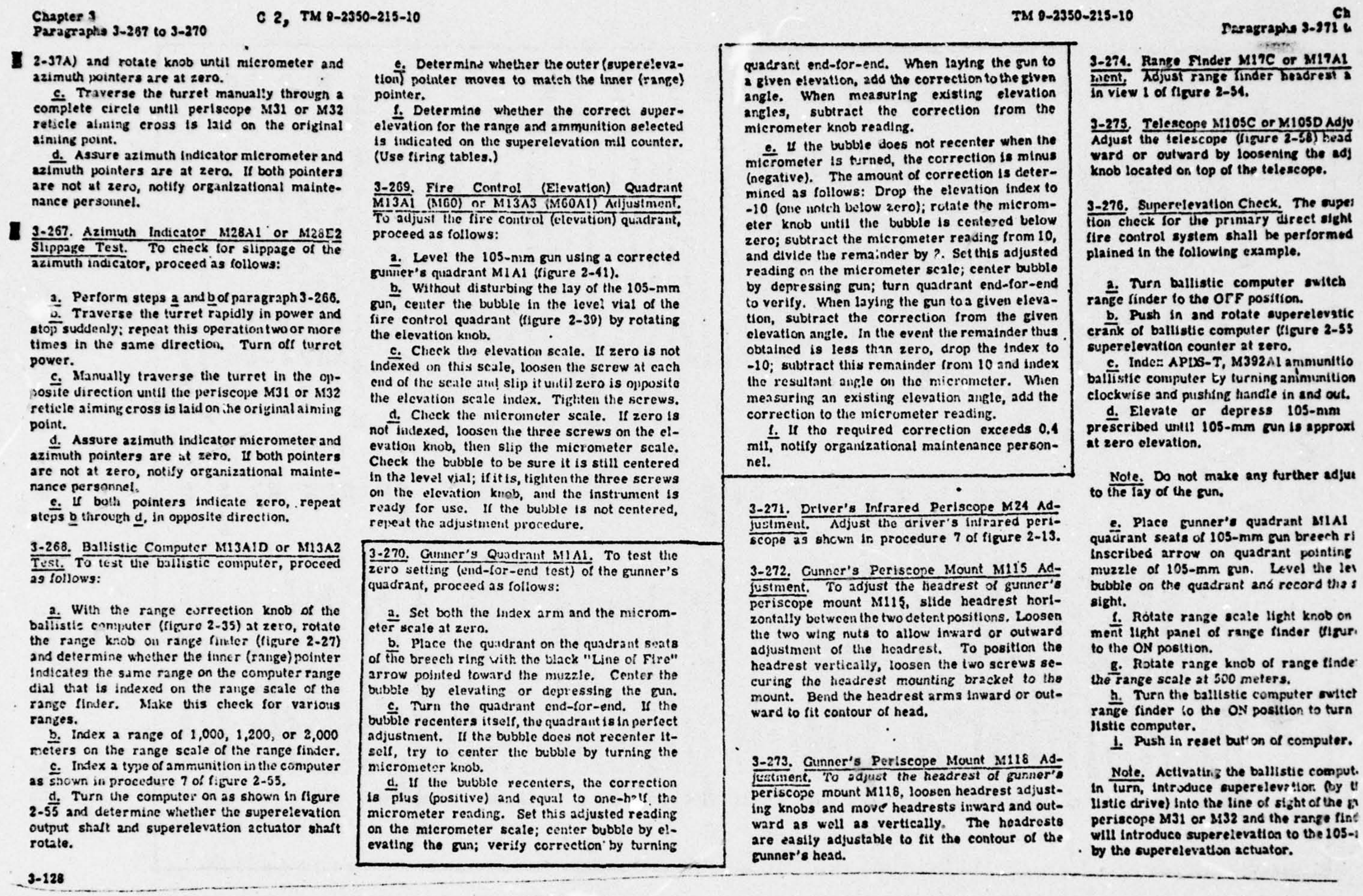
THE TANK GUNNER'S GUIDE

(Tank, 90-mm Gun, M48A1)

Figure 6-A. A Table of Contents — "Before".

TABLE OF CONTENTS

Figure 6-B. The Same Table of Contents — "After".



Chapter 3 C 2, TM 9-2350-215-10
Paragraphs 3-267 to 3-270

2-37A) and rotate knob until micrometer and azimuth pointers are at zero.

c. Traverse the turret manually through a complete circle until periscope M31 or M32 reticle aiming cross is laid on the original aiming point.

d. Assure azimuth indicator micrometer and azimuth pointers are at zero. If both pointers are not at zero, notify organizational maintenance personnel.

3-267. Azimuth Indicator M28A1 or M28E2 Slippage Test. To check for slippage of the azimuth indicator, proceed as follows:

a. Perform steps a and b of paragraph 3-266.

b. Traverse the turret rapidly in power and stop suddenly; repeat this operation two or more times in the same direction. Turn off turret power.

c. Manually traverse the turret in the opposite direction until the periscope M31 or M32 reticle aiming cross is laid on the original aiming point.

d. Assure azimuth indicator micrometer and azimuth pointers are at zero. If both pointers are not at zero, notify organizational maintenance personnel.

e. If both pointers indicate zero, repeat steps b through d, in opposite direction.

3-268. Ballistic Computer M13A1D or M13A2 Test. To test the ballistic computer, proceed as follows:

a. With the range correction knob of the ballistic computer (figure 2-35) at zero, rotate the range knob on range finder (figure 2-27) and determine whether the inner (range) pointer indicates the same range on the computer range dial that is indexed on the range scale of the range finder. Make this check for various ranges.

b. Index a range of 1,000, 1,200, or 2,000 meters on the range scale of the range finder.

c. Index a type of ammunition in the computer as shown in procedure 7 of figure 2-55.

d. Turn the computer on as shown in figure 2-55 and determine whether the superelevation output shaft and superelevation actuator shaft rotate.

e. Determine whether the outer (superelevation) pointer moves to match the inner (range) pointer.

f. Determine whether the correct superelevation for the range and ammunition selected is indicated on the superelevation mil counter. (Use firing tables.)

3-269. Fire Control (Elevation) Quadrant M13A1 (M60) or M13A3 (M60A1) Adjustment. To adjust the fire control (elevation) quadrant, proceed as follows:

a. Level the 105-mm gun using a corrected gunner's quadrant M1A1 (figure 2-41).

b. Without disturbing the lay of the 105-mm gun, center the bubble in the level vial of the fire control quadrant (figure 2-39) by rotating the elevation knob.

c. Check the elevation scale. If zero is not indexed on this scale, loosen the screw at each end of the scale and slip it until zero is opposite the elevation scale index. Tighten the screws.

d. Check the micrometer scale. If zero is not indexed, loosen the three screws on the elevation knob, then slip the micrometer scale. Check the bubble to be sure it is still centered in the level vial; if it is, tighten the three screws on the elevation knob, and the instrument is ready for use. If the bubble is not centered, repeat the adjustment procedure.

3-270. Gunner's Quadrant M1A1. To test the zero setting (end-for-end test) of the gunner's quadrant, proceed as follows:

a. Set both the index arm and the micrometer scale at zero.

b. Place the quadrant on the quadrant seats of the breech ring with the black "Line of Fire" arrow pointed toward the muzzle. Center the bubble by elevating or depressing the gun.

c. Turn the quadrant end-for-end. If the bubble recenters itself, the quadrant is in perfect adjustment. If the bubble does not recenter itself, try to center the bubble by turning the micrometer knob.

d. If the bubble recenters, the correction is plus (positive) and equal to one-half the micrometer reading. Set this adjusted reading on the micrometer scale; center bubble by elevating the gun; verify correction by turning quadrant end-for-end. When laying the gun to a given elevation, add the correction to the given angle. When measuring existing elevation angles, subtract the correction from the micrometer knob reading.

e. If the bubble does not recenter when the micrometer is turned, the correction is minus (negative). The amount of correction is determined as follows: Drop the elevation index to -10 (one notch below zero); rotate the micrometer knob until the bubble is centered below zero; subtract the micrometer reading from 10, and divide the remainder by 2. Set this adjusted reading on the micrometer scale; center bubble by depressing gun; turn quadrant end-for-end to verify. When laying the gun to a given elevation, subtract the correction from the given elevation angle. In the event the remainder thus obtained is less than zero, drop the index to -10; subtract this remainder from 10 and index the resultant angle on the micrometer. When measuring an existing elevation angle, add the correction to the micrometer reading.

f. If the required correction exceeds 0.4 mil, notify organizational maintenance personnel.

3-271. Driver's Infrared Periscope M24 Adjustment. Adjust the driver's infrared periscope as shown in procedure 7 of figure 2-13.

3-272. Gunner's Periscope Mount M115 Adjustment. To adjust the headrest of gunner's periscope mount M115, slide headrest horizontally between the two detent positions. Loosen the two wing nuts to allow inward or outward adjustment of the headrest. To position the headrest vertically, loosen the two screws securing the headrest mounting bracket to the mount. Bend the headrest arms inward or outward to fit contour of head.

3-273. Gunner's Periscope Mount M118 Adjustment. To adjust the headrest of gunner's periscope mount M118, loosen headrest adjusting knobs and move headrests inward and outward as well as vertically. The headrests are easily adjustable to fit the contour of the gunner's head.

3-128

TM 9-2350-215-10 Ch
Paragraphs 3-271 t

3-274. Range Finder M17C or M17A1 ...ment. Adjust range finder headrest a... in view 1 of figure 2-54.

3-275. Telescope M105C or M105D Adju... Adjust the telescope (figure 2-58) head... ward or outward by loosening the adj... knob located on top of the telescope.

3-276. Superelevation Check. The supe... tion check for the primary direct sight fire control system shall be performed ...plained in the following example.

a. Turn ballistic computer switch ... range finder to the OFF position.

b. Push in and rotate superelevatic... crank of ballistic computer (figure 2-55 ... superelevation counter at zero.

c. Index APDS-T, M392A1 ammunitio... ballistic computer by turning ammunition ... clockwise and pushing handle in and out.

d. Elevate or depress 105-mm ... prescribed until 105-mm gun is approxi... at zero elevation.

Note. Do not make any further adju... to the lay of the gun.

e. Place gunner's quadrant M1A1 ... quadrant seats of 105-mm gun breech ri... inscribed arrow on quadrant pointing ... muzzle of 105-mm gun. Level the le... bubble on the quadrant and record the ... sight.

f. Rotate range scale light knob on ... ment light panel of range finder (figur... to the ON position.

g. Rotate range knob of range finde... the range scale at 500 meters.

h. Turn the ballistic computer switc... range finder to the ON position to turn ... listic computer.

i. Push in reset button of computer.

Note. Activating the ballistic comput... in turn, introduce superelevation (by t... listic drive) into the line of sight of the g... periscope M31 or M32 and the range fin... will introduce superelevation to the 105-... by the superelevation actuator.

Figure 7-A. Sample of Printed Narrative.

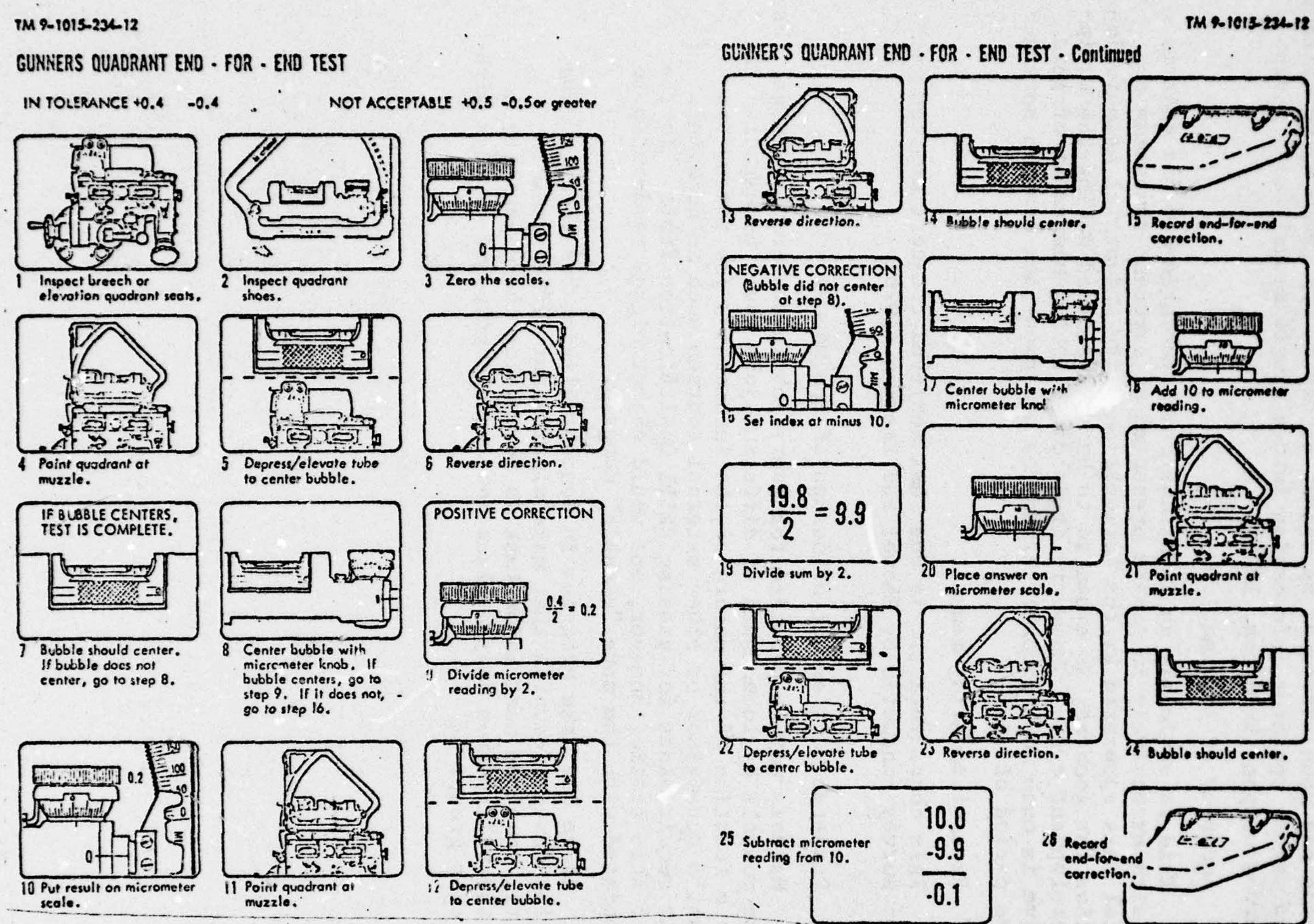


Figure 7-B. Annotated Illustrations Derived from the Printed Narrative.

information scattered throughout a given manual or across several manuals. This issue has been addressed by recent attempts, for example, to collect all crew member operator information into a single manual for a given crew-served weapon (Braddock, 1975).

7. Accuracy of information.

This is obviously an important characteristic and one that is dealt with extensively in papers presented at a recent symposium on technical data sponsored by the National Security Industrial Association. Validation for accuracy is expected to be accomplished during the equipment development cycle. However, it is clear that achievement of this objective is much more probable after the equipment has been in service for some period of time.

8. Ease of updating manual.

This characteristic is one that achieved frequent mention in the user survey conducted by Showel and Brennan (1976).

9. Readership or extent of useage of the manual.

Manuals may be provided for every job in the Army. However, use of these manuals by incumbents of different jobs would be expected to vary as a function of the nature of the incumbent's job tasks, the incumbent's supervisor or other external sources such as Department of the Army requirements for passing Skill Qualification Tests. In the midst of this static, however, one would expect higher readership to be associated with the more "useable" manuals.

If we can make progress in conceptualizing "useability" and identifying design features that make significant differences in useability, then the next major task is to develop the capability to predict useability from design plans for manuals proposed for use with a given user group.

COMMENTS ON THE PAPER BY KERN

Michael Macdonald-Ross
The Open University

Macdonald-Ross made many of his comments while critiqueing figures Kern had presented from a guidebook for Army writers. He made several points regarding design features of texts: headings, white space, and typography are supposed to be used as signposts telling the eye what is important to look at, and facilitating the eye's examination of the page; headings are supposed to tell the mind what the content is. Improper use of these design features leads to confusion. He pointed out that numbers mean order, though they are frequently used simply to itemize or to separate ideas. The latter uses can lead to a proliferation of numbers and again lead to confusion in reading. A second major point made was that expert know-how exists in the form of design specialists, expert editors, typographers, table setters, and illustrators. Much of the knowledge these expert craftspersons have is implicit or tacit, and he suggested that researchers should study these people and articulate their know-how in an explicit manner so that better guidebooks and training programs for text producers might be developed.

TECHNICAL ORDER USEABILITY

Robert C. Johnson

Air Force Human Resources Laboratory

The useability factor has not always been a prime consideration in the development of technical data. Until recent years, the basic requirement for technical data was to describe the system or component. The maintenance requirements were identified, and to varying extents some detailed information was given on how to actually accomplish the job. The traditional Air Force Technical Order (TO) is a large, bulky, hard-bound book full of words and some pictures. It does contain an enormous amount of information. Much of the time, however, the necessary data is not easily retrievable. All too often the data is not easily understood, or complete, when it is located. Specific instructions on how to do a job are often virtually lost in the surrounding paragraphs of system theory and description. Because of their large size and the bulky binder they are placed in, most TOs will not stay flat and open at the work site, and are generally too big and heavy to encourage much use on the flight line.

In the past several years, the Air Force Human Resources Laboratory's Advanced System Division (AFHRL/AS) and many other organizations have been working to improve the content and useability of technical data. One of the most popular new data concepts is generally referred to as Job Performance Aids (JPAs).

At AFHRL our work on useability has been part of a larger effort to develop JPAs. JPAs are a very special kind of TO. Having been designed and developed specifically to increase effectiveness and useability, JPAs have several very distinct characteristics. The most obvious characteristics are the small, pocket size, the step-by-step detailed procedures, and the extensive use of good illustrations (See Figure 1). AFHRL work in JPA research over the past 12 years has led to other developments that are not as obvious to the user, however. Many of these less obvious developments actually provide the real strength of JPAs.

HOW TO DEVELOP USEABLE DATA

The following processes[1] are required by our draft specification for JPAs, contained in AFHRL-TR-73-43. We believe that they are essential to the development of complete, accurate, and useable JPAs. Many of these processes would also be applicable to the development of traditional TOs.

---

[1] Specific details on the content and utilization of these processes are contained in AFHRL-TR-73-43 (Joyce, et al., 1973a). The AD numbers are AD775702 and AD775706, or they may be obtained from: Mr. Robert C. Johnson, AFHRL/ASR, Wright-Patterson AFB, OH 45433, Autovon phone 785-2606.

T.O. 1C-141A-2-4JG-4

REMOVE AND INSTALL FUEL PUMP

Install Fuel Pump On Engine:

CAUTION

Do not use any lubricant, other than engine oil on mating splines (1). Use of grease can clog oil feed holes in engine gearbox and restrict oil flow during engine operation.

6. Lubricate fuel pump mating splines (1) with engine oil, MIL-L-7808.

CAUTION

Do not allow weight of fuel pump to rest on shaft splines (1) during installation. Undue weight can impose stress on gearbox shaft oil seal, causing seal damage.

7. Working with assistant, raise fuel pump into position. Align pinholes (4) in pump with pins (3) on gearbox mounting flange (2).

8. Install fuel pump on gearbox mounting flange (2).

T.O. 1C-141A-2-4JG-4

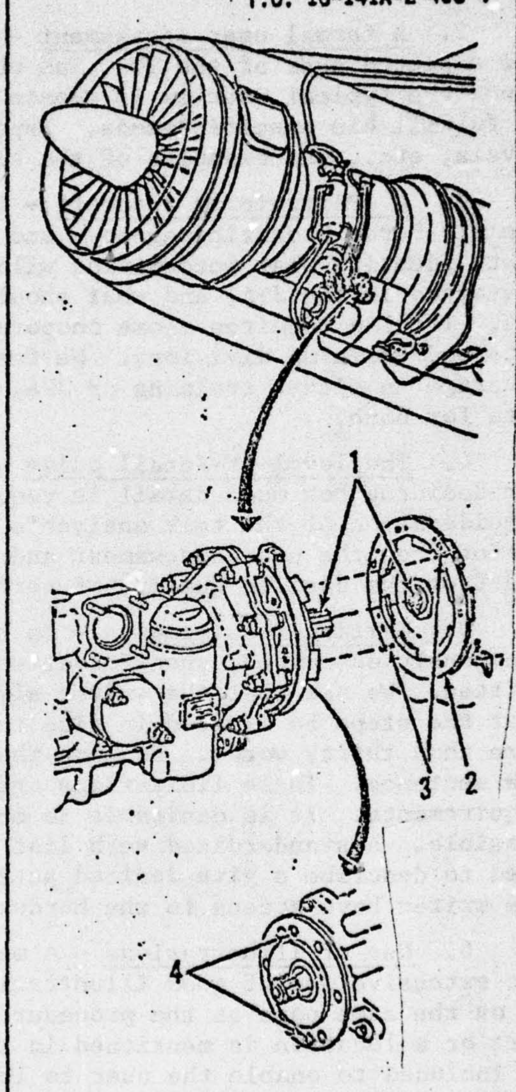


Figure 1. Example of Job Performance Aid Format and Use of Illustration.

1. A maintenance task analysis - This process forces a thorough, accurate, systematic look at the tasks required. Several intermediate products are required to insure that all tasks are identified and analyzed.

2. A formal user assessment - This process identifies and describes the expected user of the TO. The object is to gather as much information about the typical user as is reasonably possible in order to write the TO to fulfill his specific needs. Experience, training, aptitudes, reading levels, etc., are examples of the kind of data collected.

3. A JPA/training tradeoff - This process involves the JPA contractor, the Air Force procuring agency, and the Air Training Command. The object is to identify what information will be taught in training, what will be contained in the JPA, and what should be covered in both training and the JPA. It also requires close cooperation between the contractor's technical data and training divisions. We feel that this helps ensure complete data coverage in either training or JPA, and results in more realistic, accurate data for both.

4. The level-of-detail guide - This process is a set of guidelines for deciding how much detail is required. These rules are developed in consideration of the task analyst's knowledge of the system, from information from the user assessment and the JPA/training tradeoff processes, and from the desired quality of performance.

5. Writing requirements - In addition to the level-of-detail guide previously described, specific attention is given to how procedures are written. We ask that the writer always consider the user. We recommend that the steps be limited in size to no more than three sentences and no more than thirty words. No more than fifteen words are permitted in any one sentence. These limitations are guidelines, rather than absolute requirements. It is desirable to observe these limitations whenever possible. A standardized verb list ensures that the same verb is always used to describe a give desired action. It is also highly recommended that the writer have access to the hardware so he can check out each procedure.

6. Use of illustrations - A major factor in the success of JPAs is the extensive use of good illustrations. We recommend that an illustration be on the same page as the procedure, or on the facing page. Each time a part or a location is mentioned in the procedure the reference number must be included to enable the user to locate the reference on the illustration.

The illustration should be a line drawing, and it should fully support the procedure by identifying each part mentioned in the procedure. When required, it should locate the illustrated components on the next higher assembly. Care should be exercised to prevent overcrowding the illustration, and to avoid presenting excessive detail.

7. Quality assurance - Specific quality assurance provisions are required to ensure the required quality level. Included are reviews during the development process and actual tryout of data first by contractors, then by the Air Force.

8. Intended use - The writer should determine specific use requirements for the manual. Items of interest include where it is to be used, how it is to be used, what activities it is expected to support, and any special demands likely to be made on the manual.

C-141 JOB GUIDE TECHNICAL ORDERS

The first large-scale use of the JPA concept in an operational environment is that of the job guides developed for the C-141 aircraft by Westinghouse Corporation. These job guides were developed according to MIL-M-38800A. The using command, Military Airlift Command (MAC), and the contractor, Westinghouse Corporation, both agreed that a task analysis was required. These TOs, then, represent a compromise between the rather relaxed MIL-M-38800A specification and the more specific, controlled development procedures favored by AFHRL and described in AFHRL-TR-73-43.

The C-141 guides cover on-equipment work only, and incorporate logic trees for flight line troubleshooting. These TOs are in the 4 x 8" size for everything except the troubleshooting logic trees, which are contained in the traditional 8-1/4 x 10-3/4" size books.

These TOs are currently being service-tested at Charleston and Norton AFBs. This 90-day test is preliminary to Military Air Command (MAC)-wide introduction in January 1976.

C-141 JOB GUIDE EVALUATION

At the request of HQ USAF, a team has been formed to evaluate the C-141 job guides during the service test period, and for the first 10 months of their use MAC-wide. This team consists of representatives from MAC, from the procuring agency, Air Force Logistics Command (AFLC), from the Air Training Command (ATC), and from the Air Force Human Resources Laboratory (AFHRL). AFHRL responsibilities in the evaluation include user acceptance and training. Our study of user acceptance includes both attitudes and useability. Both will be described in the following paragraphs.

Attitudes - Our primary interest is to determine if any major shift in attitudes occurs as a result of greater experience with the job guides. To establish a base line we collected data at Charleston and Norton AFBs *before* job guides were implemented. We used a questionnaire that explained and illustrated a job guide, then asked questions to determine their attitude toward proceduralized data. Our sample included 124 maintenance personnel at each base. The sample was limited to the job skills of personnel who would be using job guides, but was not controlled by skill level, rank, time in service, or shift worked. Every effort was made to get as representative a sample as was possible without unduly interrupting the maintenance activity. We will again gather attitude data at both bases in late November or early December when the service test will end. In addition, some attitude data was collected by

interview during the initial implementation period at each base. We want to determine if there is a major change in attitude after repeated exposure to job guides. These data should be analyzed and available for discussion early in 1976.

Useability - Our goal in the useability study is to determine what positive and negative aspects of the job guides surface as a result of their daily use in an operational environment.

USEABILITY CHARACTERISTICS

In its most basic concept, we think of useability as the factors or characteristics of technical data which encourage their use by the intended user in the intended environment. We feel that to be useable, a TO must have the following characteristics:

1. It must be technically correct.
2. It must be easily understood by the user.
3. It must include all necessary data.
4. The required data must be easy to find.
5. The data must meet the needs of the user.
6. It must be easy to carry and use.
7. It must be readily available.

Each of these topics will be discussed and the relationships with the previously described development processes will be explained. The procedures we use to collect the data during the service test of the job guides are also described.

Technically Correct - TO information must be correct. Today's Air Force requires that TOs be used on the job. The TO is assumed to be correct. Proceduralized, step-by-step data requires virtually 100% accuracy, because any errors become very noticeable in the detailed procedures. Because the user does, or is expected to, follow every step, any errors quickly erode his faith in the validity of the manuals. Faith in the validity of the TO is necessary if the TO is to be actually used on the job. We specify the use of a task analysis, strong quality control, in-process reviews, access to the hardware, and 100% hands-on validation and verification to ensure the technical accuracy of JPAs. In the C-141 Job Guide Evaluation, AFHRL is not involved in evaluating the technical accuracy of the data. MAC and AFLC are responsible for this aspect. We do of course note and report all technical discrepancies that we discover.

Easily Understood - Technical data that cannot be understood is of little value to the user. Readability levels are quite important here. We have, however, taken a slightly different approach, and also emphasize the value of the illustration. We require a very close relationship between the proceduralized step and the accompanying illustration on the

same or facing page. Each time a part or location is mentioned in the procedure, it must be identified on the illustration. The callout number is enclosed in parenthesis in the procedure to identify and help locate the part or location. The first time a part or location is mentioned in the procedure, a location illustration is provided to pinpoint the location on the aircraft or on the next higher assembly. This helps ensure that the user can locate that part or area. In an attempt to increase the understanding of the user, we have developed a standardized verb list (See Table 1). Use of this list ensures the same verb will be used each time that particular action is desired. A preliminary information page is a part of every task (See Figure 2). This page contains all of the information the technician requires to start the job, including the special tools required, necessary manpower, replaceable parts and consumable supplies, and a brief statement about the task. Because the level of detail of the procedures also affects the understanding of the user, we suggest taking a dual-level approach. In this approach, very detailed step-by-step data is provided for those who need it, and the outline or checklist level of data is highlighted for use by the more experienced technician.

We are evaluating this aspect of useability. By observing the job in progress, we identify areas of confusion or misunderstanding. We also interview the technician immediately after the job is completed. Areas of misunderstanding are identified, and the reasons the confusion occurred are determined.

All Data Included - All of the data required to do the job should be included in the procedures. As much as possible, the book should be complete within itself. Torque values, instrument readings, pressure and electrical values expected at any given point should be provided. Referencing to other manuals is discouraged, and reference within the TO should be held to a minimum. The intention is to make the TO complete enough that the user doesn't have to refer to any outside source of information to complete the job.

Observation of the job in progress permits us to identify instances when the technician searches for information in other TOs or from other people.

Easy to Find Data - The organization or arrangement of data within a TO is an important factor in useability. A study by Losee, Allen, Stroud, and Ver Hulst, (1962) for the Aerospace Medical Research Laboratories found that 78% of the maintenance people questioned estimated that they spent more than 10% of the total job time seeking information in the TOs in order to work on an unfamiliar piece of equipment. We at AFHRL just recently replicated a portion of that study. The data are still being analyzed, but a preliminary review indicates that the percentage of time spent in searching for information is probably no better now than it was thirteen years ago. The task of finding appropriate information in most TOs can be a frustrating one. There are several obvious reasons for this. (1) The information is presented or organized in a pattern that may be logical to the writer, but not to the user. The writer often thinks and

TABLE 1. SAMPLE OF A STANDARDIZED VERB LIST

VERBS	DEFINITIONS	EXAMPLES	PREF. RANK	SYNONYMS BY ORDER OF PREFERENCE	NOTES
Accomplish	To do, carry out or bring about; to reach an objective.	Accomplish a periodic inspection on the landing gear.	2	1. Perform 3. Effect	
Actuate	To put into mechanical motion or action; to move to action.	Actuate the handpump until the pressure gage indicates 3000 psi.	-		
Adapt	To make fit a new situation or use, often by modifying.	Use the bushing to adapt the fuse to the projectile.	-		
Add	To put more in.	Add water to the battery.	-		
Adjust	1. To bring to a specified position or state.	1. Adjust the micrometer to the given measurements.	-		
	2. To bring to a more satisfactory state; to manipulate controls, levers, linkages, etc., to return equipment from an out-of-tolerance condition to an in-tolerance condition.	2. Adjust cable tension using the turnbuckles.			
Advise	To give information or notice to.	Advise man B that the brakes have been set.	4	1. Report to 2. Inform 3. Notify 5. Communicate to	
Advance	To move forward; to move ahead.	Advance the throttle.	-		
Agitate	To move with a jerky, quick or violent action.	Agitate the container so that the paint will be well mixed.	2	1. Shake	
Aid	To give help or support to; to assist.	Aid man B to lift the load.	3	1. Assist 2. Help	

JPA-31P2-2TPB1-2-3

REMOVE AND INSTALL RECEIVER-TRANSMITTER

SECTION I

PRELIMINARY INFORMATION

Introduction

This activity covers the complete removal from and installation on the AN/TPB-1A of the Receiver-Transmitter. It provides access to the Elevation Drive, the Elevation Follower, the Yoke Slip Ring Assembly, and the Azimuth Drive. Install R-T is part of the site set-up procedure. Remove R-T is part of the site disassembly. The Remove activity is always preceeded by Remove Antenna, Volume 2-6, page 5-1. The Install activity is always preceeded by Install Yoke and Elevation Drive, Volume 2-5, page 3-1.

Applicable Serial Numbers:

All

Special Tools and Test Equipment:

None

Supplies:

Sealant, Loctite MIL-S-22473
AFTO Form 349
AFTO From 350

Personnel Required:

One Technician, 30352
Three Assistants

Equipment Conditions:

Antenna removed, Volume 2-6, page 5-1. Azimuth interlock ungaged. Elevation interlock engaged at +90 degrees. Receiver-Transmitter bottom cover removed.

ACTIVITY INDEX

Procedure	Page
Remove Receiver-Transmitter	1-2
Install Receiver-Transmitter	1-10

REPLACEMENT PARTS

9 O-rings, 5A401
2 Consil Gaskets, 5A4G1, and G2
4 Connector Gaskets, 5A4G3-G6
5 Waveguide Gaskets, 5A4G7-G11
2 Seals, 5A4SE1 and SE2

1-1.

Figure 2. Example of Format for Preliminary Information Page.

writes in terms of a good working system, and successfully describes it. The user, however, nearly always approaches TOs from the view of a non-working system. (2) The actual format of the TO often hinders a data search. The block text style requires reading virtually every word of a paragraph identified by the index. (3) The writing style is often so cluttered with words that the clear, concise answer the user needs is hidden. Thus, the writing style coupled with the user's reading capability and impatience often hinders the data search. (4) The index is incomplete or inaccurate. This makes it difficult to even locate the approximate area without paging through the book.

These areas deserve some attention from TO researchers. At AFHRL, we have tried to address these areas in our JPAs. We ask that a task analysis be done, and that the viewpoint of the user be considered throughout. The task analysis forces a thorough, logical approach to data development. In arranging the data into volumes, we require that the data be arranged by sub-system, and that, in general, the tasks be in probable or most likely sequence of occurrence. More work needs to be done in the organization or arrangement of information within a book.

The format of JPAs makes it relatively easy to locate specific information. Because the data is broken down into short steps, the desired data stands out clearly after the index has referred the user to the appropriate section.

The JPA requires the developer to limit the size of the steps. As a general guideline, no step should contain more than three sentences, and no more than thirty words. No sentence should contain more than fifteen words. In addition, to decrease misunderstanding a standardized verb list is used, as is a specially developed noun list. We feel that these steps help reduce the wordiness common to many TOs.

The index is another area that deserves specific research. Our only effort here has been to use the task analysis data as a basis for the index, and thus make it more complete. We also require the common, every-day names of an item be included in the index, along with the formal engineering name of the part. We feel that the index should be alphabetical by part, not by system or sub-system (See Figure 3). Special Emphasis is placed on the index during validation and verification.

Observation and interviews are used to evaluate this area of interest. We observe the technician as he first tries to locate the proper procedure in the TO. Interviews are used to determine how easy it is to find specific bits of information in the job guides.

<u>Match the Needs of the User</u> - The needs of the user must be considered during the development of the technical data. We approach this problem by requiring a task analysis that identifies every task to be accomplished at a given level of maintenance. A JPA/training trade-off identifies which data will be taught in training courses, which will be contained in the TO, and which will be covered in both training and the TO. A user assessment process provides data which describes the intended typical users of the data. This user assessment data provides valuable information to the

JPA-31P2-2TPB1-2-3

EQUIPMENT ALPHABETICAL INDEX

	Volume	Page
Air Cleaner, 13206E0250		
Service	2-2	3-1
Antenna, Radar, 7011-7100-1		
Install on R/T	2-1	2-1
Install in Transit Frame	2-5	5-1
Remove from R/T	2-1	2-1
Remove from Transit Frame	2-5	4-1
Azimuth Drive, 7011-7500-1		
Assemble	2-6	1-1
Checkout/Troubleshoot	2-32	1-1
Disassemble	2-6	1-1
Repair	2-6	3-1
Battery, MS75047-2		
Charge	2-1	1-1
Check Electrolyte	2-1	4-1
Service	2-1	3-1
Borescope Assembly, 7011-7800-1		
Clean	2-6	5-1
Install	2-6	4-1
Remove	2-6	4-1
Condenser Pressure Switch, 1463-94		
Adjust	2-8	3-1
Connectors, 193027J3		
Repair	2-70	8-1
Crankcase Pressure Regulator, ERY 2600		
Adjust	2-1	7-1
Crimp-On Contacts, P/N 91038-3, P/N 91042-1		
Repair - See Connectors, 193027J3 Repair		

Figure 3. Format Sample for Equipment Alphabetical Listing.

JPA/training trade-off and to the level-of-detail guide. The level-of-detail guide attempts to provide the proper amount of detail for each task, according to the needs of the expected user. Some specific information provided by these processes include what a typical user can be expected to know, and what he should not be expected to know; what tools and test equipment he can use; the reading level, aptitude and intelligence scores for the typical user; and general military and civilian experience applicable to the maintenance tasks. These data allow the many needs of the user to be met, whether in reading level, additional detail to overcome lack of experience, specific illustrations, or tool and test equipment instructions.

We use questionnaires and/or interviews to determine whether the needs of the user have been met. We interview the user immediately after he completes a job. If possible, we also have him complete a questionnaire. We try to do this before he returns to his shop area and gets involved in other tasks.

Easy to Carry and Use - The physical characteristics of a manual are very important. Many of today's manuals are large, heavy, bulky books. Often the user must carry one or more manuals, his tool bag, and whatever supplies or test equipment he needs. Once on the job, the manual is too large to lay down in some areas close to the job, and in most cases, the manual won't stay open to the proper page. Illustrations and wiring diagrams are often printed on foldouts, which flap and blow, and eventually tear in the wind. The paper on which TOs are normally printed is not impervious to oil, water, and dirt, and if used, soon becomes dirty and hard to read. Constant use causes the pages to tear out of the binder. Lost or missing pages can invalidate procedures and thus increase the possibility of an accident or at least a Quality Control discrepancy.

JPAs provide some answers to these problems. First, the 4 x 8" size of the manual used on the flight line is a very convenient size to carry and use. JPAs are also bound in a binder that will lay flat and open, and it is small enough to stay with the user, even in tight quarters. Foldouts, while permitted, are discouraged in JPAs. We also encourage the using command to demand the special paper that is impervious to oil, water, and dirt, and won't tear, especially for flight manuals. These things all encourage greater use by the mechanic. The smaller, more useable size is a well-liked feature of JPAs.

We gather data on this subject primarily by observation. Our questionnaire also contains some questions on size, foldouts, etc.

Availability - A final requirement for a useable TO is that it be readily available when it is required for use. This factor is only partly within the control of the developing agency. A smaller size of manual is easier to carry and handle and thus is more likely to be available when needed. Other aspects of availability involve storage location, numbers of manuals, useage policy, and transportation to and from the typical storage location. Some organizations, such as the

Air Force Military Airlift Command (MAC), store the applicable maintenance TOs on board each aircraft. Organizations with fighter aircraft often carry the maintenance TOs in a line vehicle to make them readily available. Other organizations maintain the TOs in each work center and expect the user to carry them back and forth to the job location.

This data came from observing where TOs are kept and how they are obtained. Our questionnaire also contains some questions about availability of TOs.

SUMMARY

Useability is a very important aspect of technical data, but good useability characteristics will not occur by chance. The development processes described in this paper and more thoroughly discussed in AFHRL-TR-73-43 help ensure that the needs of the user are identified and met.

The evaluation criteria presented here are being used at Charleston and Norton AFBs. Thus far they have proven to be realistic expectations for proceduralized data evaluations. It is anticipated that further refinement of both the development processes and the useability characteristics will occur as a result of this initial evaluation of proceduralized data in an operational environment.

We believe that the results of the attitude and useability studies will greatly influence future procurements of proceduralized data. Continued research of the development and utilization of proceduralized data will also benefit from these studies.

COMMENTS ON THE PAPER BY JOHNSON

Ernst Rothkopf
Bell Telephone Laboratories

In his comments on Johnson's paper on the useability of documents, Rothkopf's main point was that there is a need to reconcile the archival functions of technical publications and the job support functions of such publications within the same training system. Archival documents are frequently held to be source documents for all sorts of doctrine, standard operating procedures, and the governance of work. They are thus a sort of stable reference for many users. The job document, on the other hand, is committed to a particular function, and that function may change: job families may be reorganized, systems may change, and so forth. There is thus a need for constant updating. Rothkopf saw this constant updating as a particular problem for job aids. He raised the questions of whether job performance aids ought to be produced centrally and, if so, whether updatings should also be produced centrally, or might local units produce their own updatings; he concluded that these are questions for future research on the useability of documents.

USEABILITY RESEARCH IN THE NAVY

William G. Muller

Naval Air Technical Services Facility

The Navy today is actively involved in documentation research and improvement efforts to positively affect the useability of technical manuals. This paper will, however, discuss only four of the many efforts presently ongoing within the Navy.

As background, an attempt will be made to provide an appreciation of the scope of a technical manual program within one of the five Systems Commands, i.e., the Naval Air Systems Command (NAVAIR). NAVAIR was chosen because it is the most complete, and it is the program with which the author has most familiarity. NAVAIR presently supplies technical manuals for 135 aircraft models. The active inventory contains 25,000 manuals. The total page count approximates *three million pages*, and each new system requires more and more pages. Figure 1 shows this growth over the past 30 years. Approximately 60 million dollars is expended every year to procure and update NAVAIR technical manuals.

Technical manuals are not written in-house. The bulk of major system publications are produced by the prime aircraft contractor. Such a contractor employs between 300 and 500 technical writers. In contrast to that, some manufacturers who supply support equipment have a technical writing staff of one and that one may be performing other tasks as well. In 1975 NAVAIR had 2,200 contractors, large and small, producing technical manuals.

The process of writing a technical manual is a long one. The time interval from the first step in the process until the actual publications are delivered is typically 3-4 years. Unfortunately, however, the configuration of the hardware system is constantly changing and, as a result, the documentation system is in a state of flux. As an example, on a recent aircraft procurement, the documentation process was started in 1969; preliminary manuals were delivered in 1972; and final formal manuals were delivered through 1975. To date almost 1,000 engineering changes have been approved. All of these changes in one way or another impact upon the documentation system. To compound the situation even more, there was a significant change in the media for presentation of documentation. A decision was made in or around 1969 to convert from paper to microfilm. This change was significant, and use is radically affected. Finally, there are those changes that appear and no one is quite sure how to deal with them. One such change is the presence of the all-volunteer service. With it came the potential for a lower reading ability. However, there was a change in the employment marketplace which interacted with the predicted reading levels. All of these changes create concern for those trying to build useable technical documentation.

The Naval Air System Command in 1970 realized a need for a research and development program in technical documentation. An analysis was

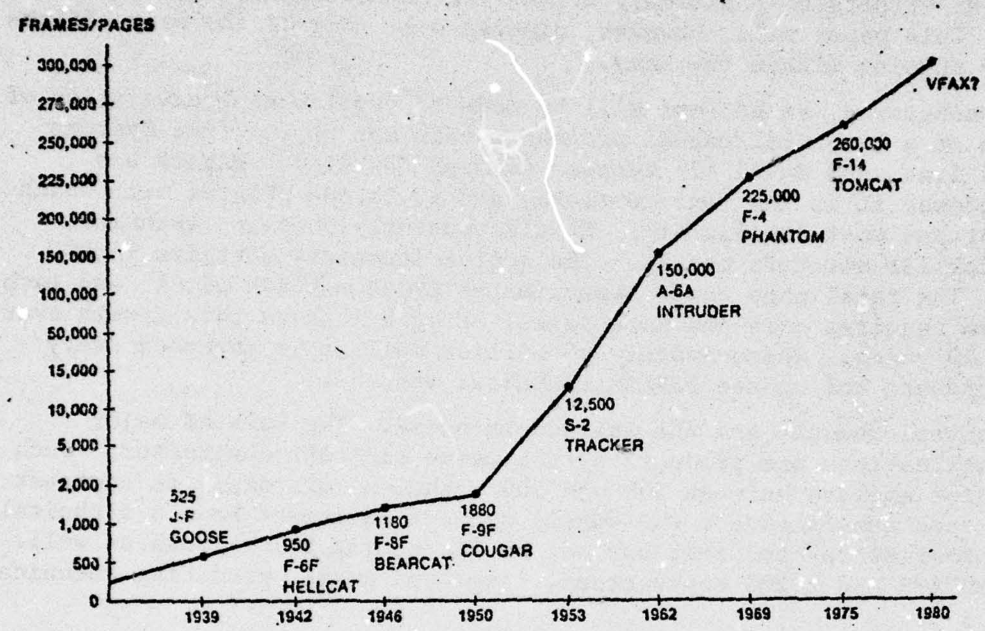


Figure 1. Technical Manual Growth Rate Within the Naval Air Systems Command.

undertaken to determine just what area in the documentation system would offer the greatest payoff. Another way of stating that is -- what are the essential elements in our system and where would available resources best be utilized?

Figure 2 is a simplistic view of the documentation system. Three key elements are shown: the technical writer, the technical manuals, and the reader/user. Outside of the technical manual area, the flow is from design to manufacturing to the delivery of the prime product, i.e., the hardware. A secondary output of design and manufacturing is a data base which serves two functions. It is first an engineering tool for design and later for modification and secondarily it is the starting point of the documentation effort. The data base is the initial input to the technical writer, who converts the data into maintenance information presented in the form of a technical manual. The technical manual is an aid to be used by the technician who is initially a reader and later a user, all in the support and maintenance of the hardware. But what is the locus of our technical manual problems? Each of these major components was examined, starting with the user and working backward.

The reader/user establishes technical manual requirements. He or she essentially sets a significant portion of the warrant for our documentation system. Reading level, comprehension skill, and anticipated training proficiency are all considerations. In addition, the work environment and the type of equipment or hardware are significant factors. Attrition is also a problem, but it is something that little can be done about directly. All of these factors must be kept in a very delicate balance. If they are not, the credibility of the documentation will suffer and when that occurs, the manual will *not* be used, or it will be used only as a last resort. Unfortunately word of mouth, recall from training, and trial-and-error experience can become the primary mode of communication when a technical manual lacks credibility. As a result, its useability can approach zero.

The technical manual as a vehicle for communication is essentially concerned with the content and format of the delivered product. What information should be in a technical manual, i.e., its content, is simple: all and only that information required to perform the task. The decision on what information to include is made by the technical writer, who will be addressed shortly. A process called validation and verification is employed to insure that the content is accurate, but this occurs *after* the manual is written. In terms of the format of the technical manual, the available approaches are innumerable. There are literally hundreds of available formats. Unfortunately, there is a propensity on the part of requiring activities to select a single format and to apply it universally. The reason for this is understandable as there do not exist today the tools or techniques to select and differentially apply formats. There has been an enormous amount of research and evaluation into format. Typically, a new format is tested against a "conventional" approach and the differences, when not masked by experimental error, are mixed. The essential quality, or that specific aspect of the format that could lead to a prescription for the format, is

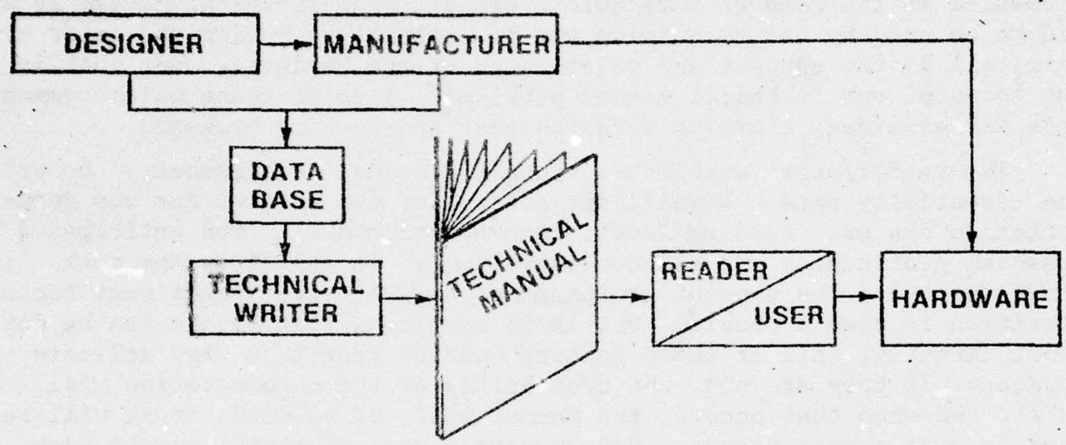


Figure 2. Technical Documentation Block Diagram.

seldom teased out of the data. As a result there has often been a go or no-go decision to apply a particular format across the board. Some recent efforts are being undertaken to alleviate this all-or-none approach. They will be presented later in this paper. An additional factor that must be addressed is readability requirements. NAVAIR advocates the use of graphics in either a complementary or redundant method to facilitate the reading and comprehension of the material. Microfilm, when used, dictates "work packaging" which presents all the material required within the confines of a limited number of pages or frames and with minimum referencing.

The last element in documentation production is the technical writer. It is here that the single largest contribution to useability can be realized. To this point, we have been dealing with the recipient of information and the vehicle of communicating the information. NAVAIR believes and operates under the principle that in communicating information, the responsibility lies with the sender. The sender in this system is the technical writer. The life of a technical writer is a veritable nightmare. The technical writer is required to convert an everchanging data base into a technical manual in a single format. Direction is provided through a legal document known as a specification, which is typically embellished by a company style guide. A potential solution presently under development in a single writing guide. The technical writer is often just one of a hundred workers and, thereby, sees only part of the total documentation. He has varying impressions, and I emphasize *impressions*, of the reader/user and the working environment. Subjectivity must be removed from the process and supplanted by objective methodology and techniques. The ability to verify writing against actual hardware is nonexistent for the greater part of the effort. In-process measures of performance must be available to the writer.

That is in summary some background on documentation system problems. It is from this viewpoint that the requirements for the research and development efforts in useability are derived. It is obvious that what is needed are techniques and tools for the technical writer that will permit the identification of the data available and to be acquired, and that will permit selection of the proper format based on an explicit and objective characterization of the user and the conditions that affect performance. Four ongoing efforts to address these requirements will be described in the remainder of this paper.

Navy Technical Manual System Program[1]

The Navy Technical Manual System (NTMS), now in exploratory development with planned completion in the 1980 time period, is a program for the development of a system of integrated procedures and equipment to improve the utility, management, production, distribution, storage, and update of operator and

[1] The Navy Technical Manual System (NTMS) was retitled the Navy Technical Information Presentation Program (NTIPP) in May 1976.

maintenance technical data (see Sulit, 1975). The NTMS program will (1) quantitatively evaluate the capabilities and deficiencies of technical manuals (TMs) in the Navy, (2) evaluate the current and future needs for Navy TMs, and then (3) design, develop, and test an integrated Navy Technical Manual System which will provide effective and timely operator, maintenance, and training data. NTMS will also provide improved operator and maintenance performance aids for the fleet, training community, and shore establishment.

NTMS is in its second year and is presently in the process of selecting a contractor for Phase 1, which is the concept formulation stage of the project. The NTMS has as its goal the integration of the various elements of the Systems Commands' normal operating systems. This means that common functions will be standardized unless they are performed differently for a specific reason. In addition, changes in the state of the art that can yield a better deliverable product will be incorporated. NTMS has developed a working model (see Figure 3) containing the essential elements of a technical manual program. These elements will be thoroughly examined and certainly expanded as the program develops. Because the evaluation of the proposals for contractor selection is still underway, further elaboration will not be undertaken in the paper.

Writers Guide

Technological advancements in aeronautical system design have transformed the casually regarded technical manual into a desirable maintenance tool. In many cases, its use has become mandatory. In these instances, the maintenance technician can no longer make decisions on specific maintenance actions without relying on the technical manual. Although routine maintenance actions will continue to be performed without reference to the written word, new support concepts, automation, and microminiaturization of system equipment have made the technician heavily dependent on the manual during the performance of his duties. Because of this steady increase in the use of the technical manual, greater emphasis is being placed on data credibility and overall useability of documentation.

The quality of the documentation delivered to the fleet has always been of interest to those responsible for its preparation. However, the renewed interest in technical manual development has generated writing and readability studies that have shown that the existing manuals are less than interesting to read and are written at a reading level that tends to impair comprehension. Further, the volume of data in conventional-style manuals complicates data arrangement, making the desired information difficult to locate. It has become increasingly evident that simplified presentation techniques must be developed to enhance useage.

Technical writers, editors, and illustrators will require special training to assist them in understanding and complying with these new and innovative techniques and requirements. The development of a writer's guide was undertaken with the intent of assisting in the establishment of standard practices that would ensure preparation of more interesting documents and also meet accepted useability criteria. It is intended that

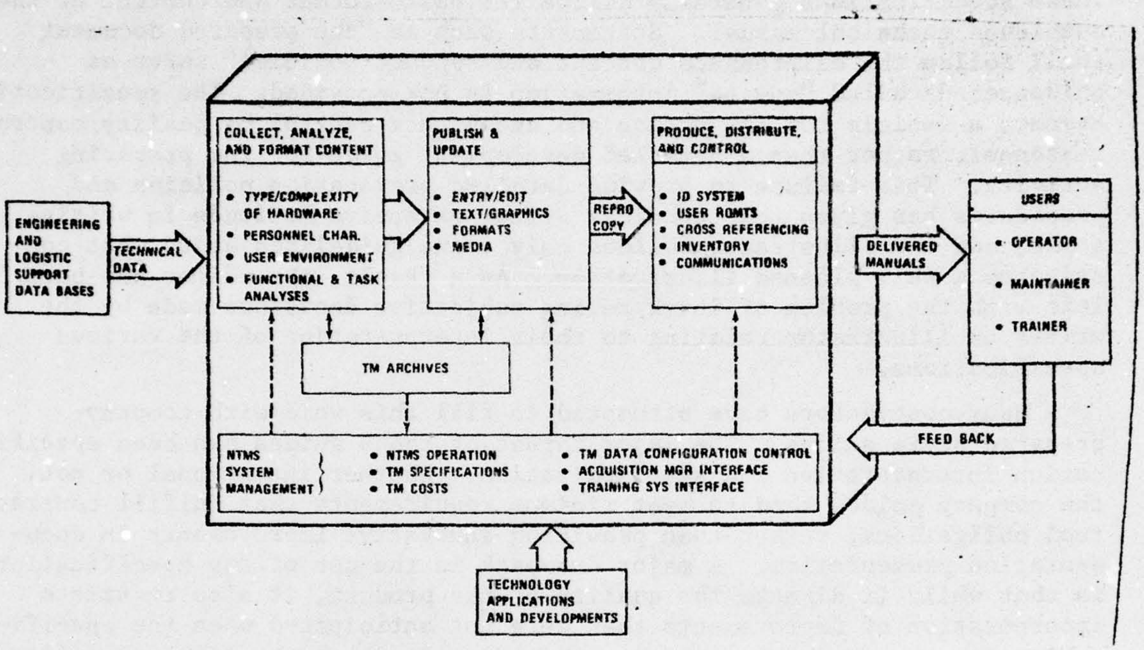


Figure 3. Navy Technical Manual System Block Diagram.

the guide be used in conjunction with the existing technical manual specifications. However, the initial issue was specifically prepared to satisfy the urgent requirements of the task-oriented "Work Package" technical manual.

Military technical manual preparation specifications provide contract definition of the two major data requirements: format and content. These specifications generally define the basic format and content of the completed technical manual. Statements such as "the prepared document shall follow the maintenance concept and support policies" serve as guidance; detailed "how to" information is not provided. The specification becomes a vehicle for inspection and acceptance control by quality control personnel, rather than a detailed development guide for the preparing activity. This failure to provide detailed preparation policies and procedures has given the technical writer extensive latitude in writing style, and the illustrator follows only vague guidelines as to what constitutes a well-planned illustration. As a result, the editor has been left with the problem of interpreting subjective decisions made by the writer or illustrator relating to their interpretation of the various specifications.

Many contractors have attempted to fill this void with company-prepared style guides. The major thrust of these guides has been specification interpretation and standardization. Whether intentional or not, the company guides tend to meet minimum requirements that fulfill contractual obligations, rather than providing innovative improvements in documentation presentation. A major drawback in the use of any specification is that while it directs the quality of the product, it also restricts incorporation of improvements that were not anticipated when the specification was prepared. Because of variations in interpretation and differences of opinion, major conflicts are evident when company style guides are reviewed and compared. In many cases, these apparent conflicts between company style guides and governing documents are the result of the contractor's attempt to respond to special requests by the government for specific types of unique publications; the style guides interpret the official guidance in relation to the specific products required.

In an effort to improve this situation, the Naval Air Systems Command has developed a writing guide (Naval Air Systems Command, 1976) to provide a single standard for preparation techniques as well as a supplement to the various specifications in force. It is intended to assist the writer, illustrator, and editor in the development of the best possible technical manual product. Task analysis, content outlines, and data arrangement are defined. Selection of logical options is explained, as are writing and illustrating techniques. Perhaps the major innovation is the incorporation of criteria affecting readability and comprehensibility requirements. The information contained has been prepared with the intent of guiding, rather than restricting, the writer's freedom of expression.

This guide was prepared to facilitate the use of separate sections by interested readers. The coverage includes these types:

- the objectives of the Work Package Concept;
- the development of a technical manual, regardless of format used;
- product analysis, task identification, and task analysis related to the planning of a technical manual;
- the development of new writing techniques;
- the preparation of comprehensive illustrations with emphasis on microform compatibility;
- techniques that will ensure credible and useable documentation, through the use of comprehensibility and readability checks.

The Work Package (WP) Manual is user-oriented. It improves on the conventional manual in three general areas: organization, format, and comprehensibility.

WP manuals are organized by functional task. The optimum WP will provide all the information required to perform a specific maintenance action in one package. This will not always be possible. However, if the information cannot be provided in one WP, it will be provided in a series of WPs. The primary task and all associated tasks that may be required to complete the job are organized in a manner that will require a minimum of searching by the technician for his information. In a conventional manual, which is organized by type of information rather than by task, it is often necessary to search through several manuals, volumes, or sections to locate the needed information.

The conventional manual, with its varied format and small type size, has provided significant problems for the microform (microfilm or microfiche) user. The format of the WP manual is designed to ensure total compatibility with either production media: paper or microfilm. Minimum type sizes have been stipulated for control of legibility. Illustrations must be properly programmed to support text. Wiring diagrams and schematics are tightly controlled. Line drawings are used in lieu of photographs.

The use of this guide with WP preparation specifications will add comprehensibility assurance criteria specifically developed for the improvement of documentation useability. The goal is to simplify the presentation of technical material and make it easier for the reader to understand. This simplistic approach has been developed through the application of rules covering writing style, use of preferred words, and comprehensive structuring of text material strongly supported by complementary figures and tables. Although developed for WP manuals, the comprehensibility assurance criteria can be applied equally well to all technical manuals.

Figure 4 is a summary of the comprehensibility assurance criteria provided in the Writer's Guide. The factors considered were not intended to be exhaustive. In fact, in order to encourage their use, there was an active attempt to minimize factors. Additionally, only those factors that could be explicitly and objectively quantified were included. The intent was to remove subjective and arbitrary items from consideration. Material such as "figure title is descriptive" or "figure level of detail adequate" were omitted due to their argumentative nature. The criteria scoring and their associated ranges were determined empirically through the use of a group of technical writers. These standards and their associated values have yet to be validated. However, the intent at this time is not to "legislate" the criteria but rather to provide to the technical writer and in-house quality assurance personnel an appreciation for factors that contribute to useable technical manuals.

Computer Simulation for Job Performance Aids

As a first step in this effort a literature search on job performance aids (Ayoub, Cole, Sakala & Smillie, 1974) was reviewed in order to obtain a workable operational definition and various categorizational schemes. The current trend of where the research is being done and what kind of research it is was also assessed. Finally, the use of computer simulation as a tool for performance aid analysis was explored.

It was found that most of the job performance aid researchers were active in the field for only a limited period of time. During the years 1950 through 1973, 78 percent of the authors in the field produced no more than a single contribution. This means that the turnover for the performance aid researchers was substantially high and thus excluded formation of any viable base of scientists for furthering the state of the art in this area.

The growth rate in the state of the art was found to be rather slow and falls far below the rates for other research disciplines. For the years 1950 through 1973 there was an increasing trend, with the peak occurring around 1969-1971. One of the main reasons for this lack of exposure lies in the fact that the performance aid literature has mostly appeared in government technical reports. This lack of circulation can be attributed to the fact that agency distribution lists are limited and not all libraries are government depositories. Only 32 percent of the literature has appeared in scientific or technical journals.

A second step in this effort was the development and testing of a computerized algorithm for the analysis and appraisal of job performance aids. The proposed approach made use of the graphical evaluation and review technique (GERT) system (Pritsker & Happ, 1966). In a feasibility study, the basic concepts and hypotheses of the approach were shown to be adequate for analyzing and assessing job performance aids.

COMPREHENSIBILITY CRITERIA	TEST	RANGE OF SCORES GOOD	FAIR	POOR
A. ORGANIZATION				
1. Access	Index/Heading	1 and greater	.5 – 1.0	0 – .5
	Index/Pages	2 and greater	1 – 2	0 – 1
	Reference/Heading	0	0 – .05	.05 and greater
2. Non-Verbal	Proportion Graphic	.50 and greater	.25 – .50	0 – .25
	Proportion Misplaced Figures	0	0 – .25	.25 – 1.0
B. COMPLEXITY				
1. Chunking	Paragraphs/Heading	2 and less	2 – 4	4 and greater
	Figures/Heading	1	.5 – 1	0 – .5
2. Proceduralization	Proportion Procedural Sentences	.90 – 1.0	.75 – .90	0 – .75
3. Consistency	Errors/Page	0 – .02	.02 – .05	.05 and greater
C. VERBAL QUALITY				
1. Readability	FOG Index	9.0 and less	9.0 – 10.0	10.0 and greater
	Flesch Reading Ease	9 and less	9 – 10	10 and greater
2. Legibility	Errors/Page	0	N/A	greater than 0
D. NON-VERBAL				
1. Pictorial	Callouts/Grid	1 – 5	5 – 7	7 and greater
	Errors/Figure	0	0 – .1	.1 and greater
2. Schematics	Components/Grid	12 and less	12 – 15	15 and greater
	Intersections/Grid	15 and less	15 – 25	25 and greater
	Errors/Figure	0	0 – .05	.05 and greater
3. Wiring Diagrams	Intersections/Grid	20 and less	20 – 35	35 and greater
	Errors/Figure	0	0 – .05	.05 and greater
4. Functional Block	Errors/Figure	0	N/A	greater than 0
5. Tables	Errors/Table	0	0 – .05	.05 and greater

Figure 4. Comprehensibility Assurance Criteria.

Using GERT-III Z (Pritsker & Sigal, 1974) and Q-GERT (Pritsker, 1974) simulation packages, a network model was developed by (1) determining or defining the basic performance steps, (2) obtaining the time required for performance of each step by an integration of experimental data and measures obtained using predetermined time and motion system (MTM), and (3) estimating the probability of error associated with the performance of each step (data obtained from human factors literature). Once the basic network was constructed, various performance conditions and limitations or resources were changed. The results yielded substantial statistical information on the effects of the various aid formats on performance time and errors (Smillie, Edsall, Sakala & Ayoub, 1975).

These results were validated in a companion experimental study which evaluated four types of performance aid formats (Sakala, 1975). The formats examined were (1) print, (2) pictorial, (3) print and pictorial combined, and (4) a hybrid group which used pictorial on easy steps and combined print and pictorial on the more difficult steps. The experimental procedure consisted of fifty location and operation tasks. The tasks consisted of movements of toggle switches, rotary switches, push buttons, pegs, and connectors. The subject viewed each instruction and then located and performed the indicated task. In addition to format, the task steps were divided in half to compare the performances of each set of twenty-five task steps; i.e., allowing the subject to become familiar with the type of instructional format as well as the apparatus. Performance differences between males and females were also examined.

Dependent measures used were time (i.e., combined time of reading the instruction plus performance of instruction), number of errors, and number of times subject reviewed task steps. The statistical analysis found less reviews for the second half of the task steps. For time, performances on the second half of the task steps were faster but interacted with format type. For the second half of the tasks the print format had the slowest time while the combined format had the fastest. For number of errors, performance on the two halves were found to interact with sex. Females had more errors than males on the first half. Females were also found to improve in performances, i.e., fewer errors on the second half.

Future embellishments of the simulation technique plan to use the statistical output to perform a cost/benefit analysis. This would allow the determination of an optimum system which incorporates the minimum of performance times and errors under prescribed conditions.

Format and Media Selection Technique

As mentioned earlier in this paper, the Navy typically chooses to apply a standard format/media technique for all the technical documentation for a given weapon system. Format in this context refers to the content and arrangement of information within a presentation. Medium refers to the means used to record, store, or display the information. Generally the display feature of the medium includes the senses the user employs to perceive the display (hearing and/or vision). The standard format/media

approach is used because of the efficiencies it allows in the production portion of the technical manual life cycle. There are minimum design costs and consistent technical manual products are attained even when a large number of vendors contribute to the overall effort. The decision on format is made early in system design. Media in general, until the advent of microform (film and fiche), had been limited to paper.

An underlying premise of this effort is that the "standardized" technical manual approach, although the most efficient for production, is not the sole factor to be considered in selecting format and/or media. An additional factor is technical manual useability.

Format and media decisions are made early in the system acquisition process. These decisions are made by Navy officials at the project management office level or by designated technical documentation representatives. Equipment component engineers representing government or contractor organizations participate in these decisions. Contracts and budgeting constraints obviously are affected by these decisions and, once made, severely limit the degree of freedom in later program stages. This effort (Post, 1975) therefore undertook the task of determining if the selection of format and media could occur before detailed design, and if, in turn, preliminary design data affords sufficient information to differentially apply format and media approaches (prescriptions) based on user data requirements (warrants). The first task was to compile and categorize the system conditions which have the potential to affect the format and media of a maintenance system.

The study isolated nineteen system conditions which have the potential to affect the format or media of a maintenance information system. These conditions, as listed below, involve equipment, personnel, and workspace considerations.

Equipment	Personnel	Workspace
Equipment Class (Types of Information)	Time to Proficiency	Illumination
Equipment Class (Access Requirements)	Personnel Turnover	Workspace
System Size	Span of Supervision	Wind
Equipment Distribution/ Layout	Navy-Wide Experience	Noise
Status Displays	Rating Selection Scores	Dirt
Support Equipment	Safety	
	Technician Response	

The study identified format and media features implied by each of the foregoing system conditions. Twenty-nine format and fourteen media features were identified.

Format Features

1. Proceduralized troubleshooting aids
2. Deductive troubleshooting aids
3. Integrated troubleshooting aids
4. System descriptions
5. Readability enhancers
6. BIT outcome index
7. Pocket-size format
8. Work package
9. Proceduralized aids for non-troubleshooting
10. Symptom/cause tables
11. SOP descriptions
12. Parts listings/identifiers
13. Job index
14. Symptom index
15. Test point index
16. Display index
17. Topic-oriented organization
18. Operator procedures
19. Debriefing guide
20. SOPS for BIT set-up
21. Proceduralized BIT usage
22. Locator pictorials by job
23. Portability
24. General purpose locators
25. Operator recording forms
26. Symptom collection aid
27. Learning features
28. Job-relevant graphics
29. Graphic density

Media Features

1. Centralized storage
2. Microform mode
3. Print mode
4. Reading portrayal
5. Reading and listening portrayal
6. Self-illuminating devices
7. Large screen portrayal
8. Head/body-mounted viewers
9. Disposable (printouts)
10. Laminated materials
11. Mini-viewers (hand-held)
12. Motion portrayal
13. Listening portrayal
14. Hooded viewers

Prescriptions, which are the application of specific format and media alternatives to a system condition, were then developed. As an example, consider the condition of personnel turnover, which may occur in either a high or low state. High turnover produces a work crew with a large percentage of inexperienced technicians. Generally, the troubleshooting aids typical of conventional manuals are too complex for inexperienced technicians. The result is that a few experienced technicians carry the troubleshooting workload. The proceduralized troubleshooting aid is an innovative format developed to alleviate this

problem. Research has shown that inexperienced technicians can perform troubleshooting when supported by the proceduralized aid form. The use of this format in this condition creates a potential for a more equitable distribution of the troubleshooting workload.

The eighteen system conditions involving 45 states and their associated format and media prescriptions were developed. The output of this effort is intended to be a guide. For ease of use the major problem or issues of presenting maintenance information were used to organize the prescriptions into related sets. Specifically, the user of the guide is asked to resolve six documentation issues by considering relevant prescriptions. The documentation issues are as follows:

(1) System description format;

(2) Recording medium;

(3) Access means;

(4) Portrayal mode;

(5) Troubleshooting format;

(6) Workplace compatibility.

The remainder of the guide discusses the six information issues and how relevant prescriptions are used to resolve them. The discussion for each issue covers the following topics:

(1) A statement of the issue;

(2) A tabular summary of the formats and media prescribed for each relevant state, and;

(3) Discussion of the rationale behind each prescription, including an illustration of the type of format or medium involved.

SUMMARY

The Navy is confronted today, more than ever before, with two cardinal facts of life. The first is that the processes used to develop technical documentation - be it a technical manual, job aids, or whatever - must be developed in a systematic and deliberate fashion. Technical documentation is at a place similar to that of technical training some three to five years ago. In training, the Systems Approach to Training (SAT) and Instruction System Design (ISD) emerged. In the near future it will be necessary to have a documentation analog, whatever it comes to be called. Three of the four efforts summarized in this report bear a direct resemblance to ISD/SAT processes. The Technical Writer's Guide calls for many of the same processes enumerated in handbooks for course developers: processes such as task identification, task analysis, and formative evaluation. The format/media selection effort is certainly seeking to create an algorithm similar to the training media selection processes contained in all ISD/SAT models. The analogy with the job aid simulation techniques is likewise obvious.

The second fact of life is that there must be continual cross communication, fertilization, and a catalyst for the integration of relevant efforts in documentation research. Certainly the NTMS program described herein has this potential. Although the primary goal of the Department of Defense is to deploy effective weapon systems and not to train better, write better or read better, the total system requires a delicate blend of both the machine and personnel components. The Navy has a strong vested interest in the communication of technical information - in communication for the purpose of training and also for the purpose of on-the-job performance.

COMMENTS ON THE PAPER BY MULLER

-George Klare

Ohio University

Regarding Muller's paper, Klare commented on the staggering effects of the large numbers and the costs involved in the Navy's technical manual system. He suggested that computer-based storage manuals, with terminals for retrieving the information, might be worthwhile to explore for the Navy. In this case, research is needed on the effects of not being able to flip backward and forward easily, and so on, on comprehension or other aspects of information processing. Klare noted an existing national library network called OCLC which currently provides access to the holdings of many libraries; plans have also been made to use a communication satellite for more efficient dissemination in the future. He suggested there may be value to the Navy in exploring this type of system for its applicability in the Navy technical manual system. Finally, Klare pointed to the need for further research to better understand and improve the useability of manuals.

PART III

OBSERVATIONS AND SUGGESTIONS

FROM THE CONFERENCE

INTRODUCTION

The *Conference on Reading and Readability Research in the Armed Services* served to bring together informed military and civilian personnel to consider the past, present, and future course of military R&D in the areas of reading, readability, and useability. The question "Where do we go from here?" underlaid the Conference discussions; suggested answers to that question emerged from the major Conference activities: (1) the presentation of perspective papers by personnel representing the Services and the response to those papers by the conferees, (2) the discussions of key issues in small Working Group sessions led by the civilian consultants, and (3) the general exchange of ideas among the conferees throughout the Conference.

Some recommendations for the general trend of future R&D in the Services have been presented in the papers in Part I of this volume. Part II contains the perspective papers and recommendations for future research and actions. The present section presents the observations and suggestions which resulted from the Working Group sessions in the Conference and from the statements delivered by the civilian Working Group leaders and by operational personnel in three Armed Services. All of these sources were, in turn, influenced by the informal interaction among conferees in response to the perspective papers that were presented by research personnel for and from each Service.

The source material for the combined observations and suggestions consisted of the following items:

1. Brief idea papers prepared by the civilian consultants, who had served as Working Group leaders. These papers presented the consultants' impressions of the Conference and their recommendations for future research in reading and readability in the Armed Services.

2. Transcripts of the recorded statements of research and development needs which the Working Group leaders reported to the general meeting on the last day of the Conference.

3. Summaries of the Working Group sessions made by the HumRRO facilitator of each group. These summaries of key issues and recommendations were prepared in part by listening to the tape recordings of the Working Group sessions.

4. Transcripts of the recorded statements of research and development needs delivered on the last day of the Conference by operational personnel from each of the three Services.

The starting point for the compilation of recommendations and suggestions was the idea papers by the civilian consultants. These papers were written after the Conference, allowing the consultants to summarize their ideas after some reflection. The editors integrated these papers into one combined paper which organized the various observations and suggestions into topic areas.

After this combined paper was prepared, the Conference transcripts and summaries of Working Group sessions were carefully examined for implications and recommendations for future research or suggested actions. These were compared with those in the combined paper. Any points which were not in the paper were added in the appropriate section. Thus, the combined paper, which begins on the following page, brings together observations and suggestions from the entire Conference.

OBSERVATIONS AND SUGGESTIONS FROM THE CONFERENCE

GENERAL RESEARCH AND DEVELOPMENT CONSIDERATIONS

Increased Coordination Among the Armed Services

There is a need for each of the Armed Services to expand the efforts to use what the others have learned. One of the results of this effort would develop a common data base and provide for regular meetings for representatives of the services to share practices, results of evaluations, and other research on their literacy programs. One mechanism for doing this would be an Interservice Committee.

It would be helpful to have some agreement on carrying out cooperative studies or at least agreement to use certain common tests for pre- and post-testing and to collect certain standard data on each of the men participating in the literacy programs.

Increased Interaction Among Civilians and Military Research and Development Groups

Both the military and the civilian research and development groups could profit from increased collaboration. University-oriented research appears to be better organized in some areas than the military's, but suffers from a lack of practical focus which the military could provide. Similarly, the military literacy programs provide ideal laboratories for instructional research. Perhaps this conference will encourage the military and civilian research and development groups to seek cooperative research and development efforts with each other.

RESEARCH ON BASIC SKILLS AND PERSONNEL CHARACTERISTICS

Understanding Marginally Literate Personnel

For developing programs that are effective, more needs to be known about the men in each of the Armed Services who need literacy instruction. This would be particularly useful for those below a fourth-grade level and also those between grade levels 4 and 8.

The causes of severe reading problems in adults could be investigated through an intensive study of 100 or so men selected randomly from those receiving literacy training in each of the Armed Services, such men to be given a thorough evaluation and diagnosis similar to one given to children and adolescents referred to a hospital or university clinic for reading and learning disability (dyslexia). This intensive study should include an evaluation of the various components of reading and related language skills (spelling, writing, auding); verbal and non-verbal intelligence (e.g., Wechsler Adult Intelligence Scale, Raven's); abilities in other basic academic skills, e.g., math; neurological and psychiatric examinations; and a case history and interview on early development, school history, and experiences.

Such studies would be of value generally, for in the civilian population more adults are presenting themselves for assistance as dyslexics. Thus, any study that might lead to a greater understanding of the cause(s) of severe reading disability among adults of normal or higher intelligence in the Armed Services would be helpful also to the general population. Indeed, studies of this kind on civilian populations selected to be roughly similar in socio-economic background to their counterparts in the Armed Services would be valuable in helping understand and cope with the literacy problem in the military. Such studies could provide each of the Armed Services with knowledge about personnel needing and getting literacy training. Specifically, they can help answer such questions as the following: What percentage of the men needing basic literacy training are essentially non-English speaking? What percent have a low ability on verbal tests? On performance tests? On both? What percent have a discrepancy of two or more years between reading achievement and mental ability? What percent are reading "up to their capacity"? What percent have medical problems—neurological, psychiatric, etc.? What percent had poor schooling? What combinations of these and other factors seem to be significant?

Such information is needed for the group and for each individual in order to provide suitable remedial programs. It is needed for knowing how best to teach them and group them, and what to expect from them in what amount of time.

Decoding Skills

The role of decoding in adult literacy instruction needs to be explored. While decoding appears to be important for children learning to read, it may not have the same value for adults who have already been exposed to initial reading instruction, but have failed to reach the level of mastery required for job tasks.

One approach to this area is to device (or borrow) a decoding test, based on made-up (synthetic) words and administer it along with a general reading test to a cross-section of poor readers. By examining the relationship between the two tests in detail, the scope of the problem might be seen.

Comprehension Skills

The question of comprehension is not a simple one. Most people would agree that to use the term "comprehension" for the full range of skills we are concerned with is to deny the complexity of the problem. On the one hand we have something we might call word meanings or understanding of vocabulary. On the other hand, we have the much more complex skills that are involved with organizing complex relationships and with search strategies. Clearly, we need more work in conceptualizing the area of comprehension.

Does a "comprehension" unique to reading exist? If not, then what has been labeled as comprehension problems in the past needs to be re-evaluated. One hypothesis is that for materials which are understood by listening, the only barrier to reading comprehension is speed of word recognition. Given a limited human processing capacity, whatever capacity required for recognizing words is not available for comprehension. This hypothesis can be—and should be—tested empirically. Speed of recognition can be measured, as can listening and reading comprehension. The relationship of speed of recognition to comprehension ability can therefore be determined for a variety of different types of materials.

One factor which this hypothesis does not consider, however, except indirectly, is reading/processing strategy. A person who has read a lot of battle descriptions, for example, usually develops an ability to organize the information which he reads in a manner that facilitates understanding. A novice, on the other hand, is usually overwhelmed by unit numbers, attack directions, equipment lists, and command structures, even if his general reading ability is high. Further attention in both research and training should be given to reading strategies—that is, to how good readers organize and recall complex information in reading.

The Automaticity of Basic Skills

Fundamental to the decoding and comprehension skills discussed above is the importance of a high degree of practice and overlearning on the skills that form the basis for higher level learning and competence. Research has pointed out that frequently a significant difference between individuals who are good and poor at reading comprehension is not necessarily skill at higher levels of understanding, but speed and accuracy in simple word decoding, a basic reading skill. The necessity of spending additional effort and using additional memory capacity on very basic skills which should be performed rapidly and automatically detracts from the ability to get meaning in the course of reading.

Furthermore, as reading passages get more complex and as basic word decoding skills are mastered, there is the additional necessity for automaticity to take place in terms of higher level units; i.e., in terms of syntactical sentence units. The inability to carry out these fundamental processes rapidly overloads an individual's cognitive capacity and detracts from the capacity to abstract meaning from a passage. The situation is not unlike expertise in a sport like tennis, where lack of mastery in the basic strokes interferes with the ability to incorporate higher order routines and strategies. The individual's capacity is overloaded with concentration on the basic skills; development of automaticity in these skills would free him for higher level performance. Research on the nature and development of automaticity in retrieving the basic concepts and using skills that underlie functional job literacy, general literacy, and institutional literacy[1] would seem to be a research and development activity with high payoff.

---

[1]These terms are discussed in the following section.

CONCEPTUALIZING THE NATURE OF LITERACY AND OTHER COMMUNICATION SKILLS

Concepts of Literacy

During the conference, an individual in the Armed Services was characterized in terms of three kinds of literacy: functional job literacy, general literacy, and institutional literacy. *Functional job literacy* pertains directly to competence and expertise on a job. It involves the knowledge of important job concepts and the relationships between them, the ability to perform complex skills and routines, and the ability to be able to use and retrieve these knowledges and skills in appropriate situations in an efficient manner. Functional job literacy can be significantly influenced by job performance aids of many different kinds.

Significant questions were raised in this area. One pertains to the relationship between job expertise and abilities in verbal literacy skills. The strength of this relationship varies for different jobs, and might there not be a tendency to "overliteracize" the requirements for particular jobs? Training may be designed so that it is more verbal than it should be, since there is an instructional bias for teaching and testing in a verbal manner. Verbal teaching carries with it elements of prestige and often is easier to carry out. A second question pertains to the recognition that job expertise involves the development of a network of concepts in the mind of a competent individual. The question raised is how various job concepts are interrelated for the expert, and how this differs from the knowledge structures of the novice or the less competent individual. Current advanced research in cognitive psychology is beginning to develop ways of displaying "conceptual semantic networks," and the techniques of this kind of research may be useful in describing how knowledge is organized for the expert so that such knowledge relationships can be duplicated through appropriate training and experience in less competent individuals. Similarly, longitudinal studies and contrastive analysis research on the functional job literacy performance of good and poor readers may yield knowledge which could be incorporated into appropriate functional job literacy training for the poor readers.

General literacy pertains essentially to the kind of literacy that is necessary in order to be able to enter training programs leading to functional job literacy. This kind of literacy involves basic intellectual skills; i.e., certain levels of competence in the basics of reading, writing, arithmetic, and fundamental cognitive skills like those measured by various aptitude and intelligence tests. The development and enhancement of general literacy skills can increase lateral mobility so that individuals are not restricted to a few job specialties and have a wider choice of available careers or school possibilities.

Two approaches to developing general literacy are possible. One approach involves the direct teaching of the skills and competencies that are to be learned in elementary school and junior high school. This is

simple enough to state, but often difficult to implement, because of various motivational factors or an individual's feeling of not being able to get started, etc. One interesting possibility for breaking down these barriers is to begin individuals in functional job training programs first so that they can learn by hands-on experience and with a minimum of verbal superstructure. This initial experience could provide individuals with some feeling of competence and the ability to carry out interesting and useful duties. Once an individual develops a basic set of job concepts and related job skills, a knowledge and motivational structure may be established that encourages and enables the individual to learn general literacy skills that can permit him to advance further in his job. In this way, the general literacy block might be broken because possibilities have been opened for which these skills seem highly relevant and important.

A second and long-term research approach to enhancing general literacy skills is to carry out basic research on the cognitive processes that underlie the nature of intelligence and aptitude. At the present time, this kind of research is not too abstract nor far-fetched. New research endeavors are investigating the processes that underlie the psychometrically measured entities of intelligence and aptitude. These entities now are being interpreted in terms of cognitive information processing skills that may be amenable to specific training.

Institutional literacy involves the information and general know-how required to get along in the military setting. This includes the ability to understand regulations and to participate successfully to those activities which contribute to one's advancement.

Literacy, "Numeracy", and "Dexteracy"

The emphasis of this Conference on reading and readability leads to a concern about the "overliteracizing" of training, even though there is excellent work going on that attempts to decrease literacy requirements in job manuals by including more pictures, diagrammatic explanations, etc. However, it needs to be reiterated that individuals can learn in different ways, and that for some individuals, initial emphasis on verbal skill may not be the avenue to further competence. The road to competence might best be approached by initially emphasizing hands-on dexterity by which individuals can develop expertise up to a point which can then incorporate necessary verbal components. So, in addition to literacy, Glaser used the word *"dexteracy"* to emphasize the fact that literacy in our society may be overemphasized to the extent that the dictionary does not contain such words as *numeracy* and *dexteracy*.

The fact that individuals learn in different ways means that research must seriously examine both the verbal and performance components of training and the particular strengths of individuals being trained, and optimize the relationships between the two. There is a nice example to illustrate this point that relates to taking advantage of the particular strengths of individuals and adjusting training accordingly. The study

(Wing & Wallach, 1971) considered the characteristics of students who apply for admission to college. Examination of the workings of an actual college admissions committee indicated that the usual procedure was to give strong emphasis to standardized Scholastic Aptitude Test scores and to high school academic achievement in judging the acceptability of applicants. Less recognition was given to outside-of-school attainments, such as science projects or special employment. The college admission process, therefore, skimmed off the cream with regard to test scores and grades but failed to grant explicit recognition to individual attainments which expressed themselves in nonacademic ways. A simulated admission procedure in which nonacademic accomplishments were considered as part of the college selection procedure resulted in the selection of an admitted class whose membership differed substantially by some 50 to 60 percent from a class admitted by the traditional practices. Furthermore, the hypothetical new class represented a wide spectrum of attainments and performance capabilities. This kind of study suggests that selection procedures of the traditional variety tend to restrict the possibility of adjusting to the performance capabilities of individuals in order to maximize success in various occupations. If this broader range of capabilities were considered, educational content and teaching techniques would have to adjust to the specific real-world accomplishments that students could bring to bear on their schooling.

While this example deals with a high level college population, the emphasis on adapting to special attainments derived from either outside the military services or in them is worth considering as one possible way of adjusting to the capabilities of individuals as a basis for their further training. By virtue of particular experiences such as hands-on apprenticeships, individuals develop expertise from which they could have been precluded by an overemphasis on "literacy" as compared to "dexteracy." An interesting research activity would be to apply the tactic employed by Wing and Wallach for a population like that entering the military services. An important aspect to consider in this regard is that people can develop higher levels of specific talent that may not be correlated with measures of general verbal literacy (IQ). Excellent chess players, mechanical and electronic experts, for example, can achieve competence through largely nonverbal routes, such as the visual modeling of others' performance and practical experience.

Measuring Literacy and Other Communication Skills

Measures need to be developed for the Armed Forces, based not only on reading, but also on writing, speaking, and listening. The listening scale is needed to determine whether a recruit can comprehend oral English sufficiently for a task; if he cannot, a literacy course would be a waste of time. This, in essence, is what the current auding-reading comparisons are aimed towards. (See Groff's paper in this volume.)

Speaking and writing are equally important for some military tasks, yet seldom receive the attention which reading does. A scale developed for listening and reading could be extended, with some modification, to

speaking and writing. Hopefully, competence in these latter two skills will be (or is) considered for promotion.

It should be noted that already a fair percentage of new recruits, especially into the Army, do not speak English as their native language. A full utilization of this manpower resource requires that oral English be assessed and taught. In building such tests, it should be pointed out that grade level norms are not satisfactory criteria for adult literacy goals; whether they are useful even at the elementary levels is debatable. Besides their arbitrariness, grade level norms are based upon vocabulary, subject matter, and syntax considered appropriate for elementary children. To apply these to adults is in part inconsistent and in part degrading. A more useful approach to setting reading criteria could be based upon the Job Reading Task Tests which have already been developed for the Army. If the goals of Armed Forces literacy programs are task-oriented, then task-oriented criteria are far more appropriate than criteria based on general vocabulary and comprehension.

LITERACY PROGRAMS AND TRAINING SYSTEM CONSIDERATIONS

The Study of Literacy Training Processes

In the presentation of training studies, much concern was given to selection variables, to the measurement of output achievement, and to the measurement of short- or long-term effects of training on job performance. Less detailed attention was given to what actually happened during training. There is obvious attention to the development of training packages and to the comparison of different forms of training. However, little study seems to be made of the details of actual training implementations. For example, when training method A is compared with training method B, the programs are described only in terms of gross components; little information is provided about the details of how they are actually implemented, and about the dimensions of each that were successful or the dimensions that were unsuccessful. It is difficult to compare brand X against brand Y primarily by their labels. It is necessary to know whether actual training conditions were better for one than the other, whether field conditions reduced one to being the same as the other, whether one involved a better group of instructors, or whether one permitted more time to study, etc.

What is required in assessing various training procedures is the detailed analysis of classroom processes and implementation conditions. This requires isolation and observation of a number of implementation processes that include such dimensions as (a) opportunity for learning - the time available for actual on-task involvement; (b) matching - the relationship between what is actually taught and what is actually tested; (c) material motivation - the degree of interest built into the training materials; (d) teaching motivation - the degree of excitement engendered by the instructor; (e) structure and sequence - the extent to which a curriculum is carefully sequenced in a form of a hierarchy of prerequisites so that there is opportunity for mastery of earlier skills before going on

to more complex skills; and (f) monitoring and feedback — the degree to which the training process allows for the detailed analysis of student strengths and weaknesses in the course of learning, and how instruction is adjusted to this information. An interesting research endeavor is the conduct of field studies that investigate what is actually taught and tested, and how it is taught, so that processes of training can be improved and so that different training systems can be given adequate comparison. A training system may also be compared to a theoretical conception of the expected change in the reader to determine whether the training has been effective.

There is need for a study of the teachers of the literacy programs to determine which competencies are necessary, how they may be developed, and how they can be evaluated.

The amount of time necessary to effect a significant improvement in reading achievement needs study. At the Conference, it was made clear that only a limited amount of time has been scheduled for literacy training, even for the men who come with no reading ability whatsoever. It is suggested that the general, rather disappointing outcomes of the usual literacy training programs may come from expecting the impossible; i.e., expecting 4 or 8 years of progress, in terms of standardized reading achievement tests, in 8 weeks or so. Even if some of the basic skills can be acquired in so short a time, it is questionable whether the integration of these skills can be acquired in this time. What is needed by children (and it is suggested that illiterate adults do not differ substantially in this respect from children with similar reading levels) is extensive reading and practice over an extended period of time. How long a time is needed for adults to reach certain reading levels should be investigated. Specifically, it should be determined whether the amount and timing of the instruction and practice varies with the entering reading level, mental ability, and the kind of reading program used to teach the beginning reading skills.

Study is needed not only on the amount of time necessary to effect significant improvement in reading achievement, but also on the amount and nature of improvement that qualifies as "significant". What does a given reading gain mean in terms of job performance? Just as grade level norms are unsatisfactory criteria for adult literacy goals, they are unsatisfactory in evaluating the effectiveness of literacy training programs.

Studies of Alternative Delivery Systems

Media

Regarding the delivery of literacy training programs, it would be well to look into the feasibility of making greater use of media - TV, films, tapes and texts, etc. While there is serious study underway for producing such teaching materials, it is suggested that the lessons learned

from such programs as Children's Television Workshop's Electric Company be examined for possible use in teaching basic reading skills to men at the lowest level of reading achievement.

Computer-aided instruction is still in an early experimental stage, so far as its capability to instruct is concerned. Nevertheless, for the Armed Services, where available time for instruction is severely limited, CAI might be beneficial for some types of instruction. It would be less sensitive to manpower change than the present instructional system, and considerably easier to update than videotape, films, or any other passive medium.

In addition, CAI might be used for continuing education on a free-choice basis, thus allowing military personnel the opportunity to take courses on their own time. The PLATO project at the University of Illinois has developed supplemental reading instruction for the primary grades. Some of their experiences might be transferable to an adult setting. The University of Wisconsin is experimenting with the use of PLATO as an on-line, interactive system for diagnosing reading skill deficits. The system has a complete test analysis subsystem built in so that test results can be printed within minutes after the last person is tested. Other applications of CAI, particularly those which provide job-functional literacy training, should be considered.

Integrated Literacy and Technical Training

All of the Armed Services share the need of having men perform very specific occupations with high degrees of proficiency. When men do not perform their jobs or cannot learn to perform their jobs, there is a problem of occupational competence. Literacy is one important aspect, but not the only factor of proficient job performance. In the Armed Services, which define their mission more in terms of specific job training than general literacy training, the extensive investment in job performance aids shows an implicit recognition that the issue is primarily one of occupational competence.

The humanistic value of literacy training may also be redefined as a need for basic competency training. To effectively transmit literacy to marginally literate men, it will be necessary to place reading in the context of a larger set of competencies. Essential basic competencies for adults include reading, writing, speaking, listening, calculating, and problem-solving. It is reasonable to suppose that these competencies are connected and that improvement in one will not be useful and durable for the individual without concomitant improvements in the others. What are the listening, speaking, reading, writing, computing, and problem-solving competencies required for a given occupation? This set of competencies may not be completely appropriate for the types of jobs that occur frequently in the Armed Services. However, it is clear that reading technical documents would be one major aspect for a large proportion of jobs. Another major aspect would probably turn out to be problem-solving. Indeed, problem-solving with written documents was mentioned frequently in this

Conference. Good job performance is associated with frequent reference to reading matters. A cause-and-effect relationship cannot be inferred here, since high intelligence could mediate between the two variables, but the relationship is suggestive. In fact, visual search of written documents seems to be a relatively frequent and relatively difficult aspect of job performance. There is, of course, a heavy technical knowledge load in many jobs.

It may be instructive to consider the analogy between literacy training in the Armed Services and literacy training in developing nations. Of course, there are points of difference, but the points of similarity are useful. First, it should be noted that, as reported in the 18th session of the General Conference of UNESCO in Paris, 1974, there is a substantial undertaking known as "The Experimental World Literacy Program and Its Global Evaluation". Over a five-year period from 1967 to 1972, the program reached more than one million people (1,028,381) in over 20 thousand (20,379) instructional classes and employed a total of over 23 thousand (23,480) instructors at a total cost of about 27 million dollars. While the programs in the countries vary, they all draw a sharp distinction between "traditional literacy" and "functional literacy". Traditional literacy programs have defined literacy as reading at a given reading grade level (i.e., reading at the 4th- to 6th-grade level) and the training programs attempt to bring adults up to that level of general reading ability. For many, varied reasons there has been a widespread disenchantment with this conceptualization of literacy. For example, 4th- to 6th-grade reading level did not provide immediate social and economic returns to people. Low motivation and inadequately prepared teachers contributed to attrition rates that were so high that teaching could not be effective. In its place, the concept of "functional literacy" has been substituted. This concept embraces reading, computation, and knowledge in a specific area of social value, usually agriculture. For example, reading, computation, and knowledge about modern cotton-growing practices are taught simultaneously in Tanzania. This type of program reflects a more complete and realistic occupational analysis than former programs and it is vastly more successful.

At the UNESCO conference mentioned previously, the following principles of literacy training were tentatively forwarded, based on experience in these projects:

1. Functional literacy programs should be geared to specific objectives, problems or centers of interest.
2. There should be an integration of various components (language, mathematics, technology, scientific acculturation, scientific training, graphic representation) into a unified educational process.
3. Programs should be adapted to socio-economic and geographical conditions.

4. Teaching should be intense enough to produce desired results in a minimum time period.

5. There should be constant adaptation of programs and teaching methods to the changing technology and demands of development.

6. Literacy training should be regarded as the first stage in a process of lifelong education.

While the populations in developing countries and the Armed Services vary in many ways, it is noteworthy that in both cases occupational competence is a major goal. In developing countries, it has been found to be imperative to train competencies such as mathematics and basic knowledge alongside the literacy skills of reading and writing if the goals of economic productivity and social advancement are to be achieved efficiently. Perhaps the same principle would apply to the Armed Services in the United States.

One of the problems with general literacy training programs is that the programs do not improve a man's reading ability to the level necessary for functional use of the reading ability in job performance. Increasing a man's skill from 4th-grade to 6th-grade reading level is valuable, but it is not functional for many jobs. In a functional literacy program, men would be trained to learn to perform specific reading tasks to the level required for adequate job performance. Job technical knowledge, concepts, and skills training would be ingegrated with literacy training including job reading tasks; e.g., reading orders, searching the manual of inventory, calculating prices, etc. Men could stay in a training course until such time as they passed a proficiency examination with criterion-referenced, job-related reading task tests. Thus, men who needed to learn a number of basic operations, such as reading and writing, would probably stay in the course longer than men who were higher in these competency areas. This suggests a deviation from some of the current practices in the Armed Services. At present, in many instances, the intention is that men will first be trained to read and will then be assigned to job skills training. This adds to the total training time and fails to capitalize on the immediate transfer to be gained from integrated, functional literacy and occupational training. One issue that is raised by such a practice is the extent of transfer of functional literacy training to general life situations. Research could be performed to explore the generalizability of job-specific literacy training.

The overall thrust of the notions expressed above, particularly those that pertain to the improvement of functional literacy and to adjusting job skills training to the developed capabilities of individuals, suggest a model of training that is different from the model generally employed in the military and in most civilian training and education. The usual training system is a *selective* system which emphasizes the selection of individuals on the basis of preliminary screening tests that predict who has

a reasonable probability of succeeding in a particular system of training. The system of training is set-up for those individuals with specific entering credentials and is improved by tightening credentials, but little adjustment is made for individuals who might profit from extensive prerequisite training in order to get through the system or by changes in the training system itself to enable them to attain the full range of competencies that are required and taught by the training system, including literacy skills. Research is needed to explore the feasibility of developing the more *adaptive* training programs needed to provide the full range of competencies personnel may need to have developed in order to successfully acquire job knowledges and skills.

Research might also explore the feasibility of extending such an adaptive training program to include pre-service functional literacy training in broad occupational areas. Policy decisions permitting, such training might be offered by the Department of Defense acting for all the Armed Services or coordinated with other governmental agencies concerned with adult basic education in job-related contexts, such as the Department of Labor and the Office of Education.

RESEARCH ON THE READABILITY & USEABILITY OF MATERIALS

Research to Improve the Predictive Efficiency of Readability Formulas

Agreement of Definitions of Terms

Such terms as "readable", "readability", "comprehension", "comprehensibility", and "useability" should be reviewed and standard definitions suggested for Armed Services useage. Standard definitions would prevent misunderstanding such as apparent differences in data which are really differences in interpretation stemming from differences in terminology. The literature on readability provides a number of examples of disagreement in interpretation that, on closer inspection, become merely confusions in useage. As a case in point, the term "high readability score" may refer to something that is easy to read or something that is hard to read. The reason is that with some formulas (e.g., the commonly used Flesch Reading Ease formula) a high score means easy reading, while with others (e.g., the commonly used Dale-Chall formula) a high score means hard reading. Examples could be provided for the other terms as well. Standard definitions might not only prevent unnecessary confusions and disagreements in Armed Services publications and useage, but might (hopefully) spread to educational publications and useage as well.

Criteria for Predicting Difficulty of Materials (Readability and Useability)

A problem discussed at the Conference was the tendency to talk about written materials as though they were all similar in structure and purpose. A useful distinction was suggested between "operational" manuals, or ones that tell how to perform job tasks, versus training manuals, which are to permit the learning of job information. Though one might wish to use a readability formula on both types of manuals, the desired criterion for operational manuals might be task performance, while the use of recall, recognition, or some other learning measures might be best for evaluating the readability (learnability) of the training manuals. With regard to operational manuals, the value of the document might be assessed by measuring the time required to perform the task with and without the document within a defined population of men (with general aptitude, p, and job experience, q). Thus, the value of the document could be determined by the ratio of the "with" and "without" performances. Such a criterion could be applied to schematics and reference manuals, as well as books or other technical writings.

Such a criterion would be useful in attempting to improve certain documents. For example, the extent to which a manual is well organized and lends itself to search operations could be determined by examining the amount of time required to locate information in the document. The quality of the document could be systematically improved by manipulating its characteristics in such a manner as to reduce the time needed for completing search tasks using it.

Regarding the use of learning measures to evaluate the readability (learnability) of training manuals, it was suggested that such measures might be an intermediate evaluation step in the case of narrative materials used for training, where understanding of principles is highly desirable, and that some kind of task performance might be a final outcome for training manuals, as well as for operational manuals. Another suggestion is to use as an intermediate evaluation step, a "levels-of-comprehension" distinction, with a deeper level (e.g., one that permits more transfer of training) desired in the case of training than in the case of operational use. Differences among the Armed Services on this question might also be explored.

In addition to possible differences in criteria to be predicted by readability formulas in the case of operational versus training manuals, criteria for "comprehensibility" and for "useability" should be reviewed and the value of the two concepts compared for Armed Services useage.

The term "useability", though uncommon in educational publications, is becoming common in Armed Services circles. Criteria for measuring useability, however, are not commonly agreed upon, and should be reviewed. The term "comprehensibility", on the other hand, is common enough, but the criteria used for measuring it give rather strikingly different results. A prime example is the use of multiple-choice versus cloze comprehension scores. The former measure typically yields higher percentage correct values than the latter, but validity coefficients for the same readability formulas typically yield higher values for the latter than the former.

Research on the Indicators of Readability

Currently available indicators of style difficulty (i.e., of the psychological demands of the expository form of written material) are not wholly satisfactory. Further investigation of the psychological basis for readability indices might be considered. Several interesting models for the processing of text have emerged and have withstood empirical tests sufficiently well to warrant exploratory work on the applications of these models to practical documentation work. Some care is indicated, however, to assure that what is done is guided not only by the basic scientific ideas, but also by a realistic determination of what can be realistically applied in practical settings. (See Curran's paper in this volume for a discussion of the difficulty of computation of a readability formula versus increase in predictive power.)

Questions arose a number of times about whether readability formulas are predictive for highly technical prose and for semi-prose material. A grid was suggested that might have kinds of writing along one axis and application situations along the other, with empirical answers based on existing data located at the intersections of the grid. Where data do not exist, it would be immediately apparent from the grid that studies were needed, *a la* the classic periodic table in chemistry. This is not to suggest that the degree of precision would match that of chemistry, but the procedure could still be useful. Although it would be desirable to have all of the empirical data based on use of the same formula, for example, a first approximation would not demand this. Similarly, though it might be desirable to have an index of the degree of technicality of writing for descriptive purposes, crude judgments might be sufficient for the first approximation. In time, if this approach proves helpful, a more complex grid and/or other grids might be developed.

Readability formulas, because of their characteristics, should be thoughtfully applied. One such characteristic is that they are content-free. Since word difficulty or word length and sentence length (or sentence complexity in some cases) are the predictors, the problem of a reader's knowledge of content must be appraised in some other way. The extensive use of technical terms and the information base an individual must have to cope with these terms cannot be taken into account through the use of formulas.

Many conferees recognized the critical importance of the content of written material. Several judged content to be a much more important indicator of the readability of written documents than expository style and form. Yet the reports read at the Conference (as well as the substance of the afternoon discussions) indicated with painful clarity that methods for the systematic assessment of content are rudimentary at present. In addition to assessment in terms of the knowledge base which a reader must cope with, the development of effective methods for determining whether the content of written material is sufficiently complete and accurate to achieve the objectives of documentation deserves high priority in future practical research sponsored by the military Services.

One approach suggested for taking content into account in readability formulas is to consider the technical terminology of manuals as a separate factor in predicting readability. This general kind of question — taking special terminology into account to improve formula predictiveness — has been around for some time. With a word-list type of formula, this can be done by adding words *to* the list or by excluding certain words *from* a count. With a word-length type of formula, this might also be done by excluding words from the length count during application, but this can cause considerable inconvenience and, in the case of human (as opposed to computer) application, probable unreliability of application. A better possibility would seem to be the use of a correction factor based upon the technicality of the material. A technicality index might be developed for this purpose, using (for example) number of technical terms in a sample of text over number of total words, or, preferably, over number of content (as opposed to structure or function) words. The usefulness of these approaches is, of course, an empirical question, and studies should be conducted even though (or perhaps because) their success is by no means assured. Trying to add another index variable to formulas or to correct existing index variables may unfortunately be less effective than we might hope; adding an adjustment for technicality of the material may produce no increase in the predictiveness of the formula. But again, the question is an empirical one and should be investigated.

Research to Improve the Production of Materials

Research on the Selection and Training of Technical Writers

While readability formulas are somewhat efficient in *predicting* the reading difficulty level of materials, they do not offer guidance for *producing* more effective, readable materials. There is a need for research to help train writers to produce more readable/useable materials. One approach suggested is to identify the better performers among groups of writers in terms of their "track record" in producing good technical text over a period of time. A *post hoc* analysis might then help eventually to answer such long-standing questions as whether it is better to teach a technically trained person to improve his or her writing skills or to teach a trained writer enough of the needed technology to permit him or her to write accurate technical documents. Even if one answer will not fit all situations, this approach might help to say something about when one or the other possible sources of writers should be considered.

If the characteristics of good technical writers can be identified, further work on the development of selection instruments should be undertaken. Discussion indicated that many companies that produce materials for the military do not keep any certain number of technical writers under employ at all times, or at least do not keep those they have at technical writing tasks all of the time. Consequently, some selection procedure could be helpful during crests in the need for technical writers. If a

test-type instrument proves difficult to develop, a profile-type approach might be considered, making use of such variables as background, training, experience, interest, etc.

For training purposes, research is needed on what to tell writers who wish to improve the effectiveness of their writing or who, alternatively, have been told their writing is inadequate and must be changed. Discussion suggested that writers seldom, if ever, set out to write less effectively than they can. Yet this happens, and most writers sincerely welcome really useful suggestions for improvement. What suggestions can be given? Writers' guides fail to agree, and in some cases, actually disagree, on what should be done in writing clearly. One approach is the Army writing guide that is soon to be published. (See Kern's paper in this volume.) Kern found that Army writers seem to "model" their documents after those that have gone before, and this guide suggests preferable alternative models. A second writers' guide produced recently by Klare presents suggestions for readable writing based upon psycholinguistic and readability research. The effectiveness of both of these guides, as well as of others, needs to be systematically evaluated. Such evaluation should include attitudes of writers toward such guides, probability of use on the job, et cetera.

Writers' guides should also deal with the treatment of graphic materials, for the work of technical writers involves decisions affecting the design of graphic as well as textual material. Suggested research and development activities include the following:

a. Research and develop guidelines for the use of graphic materials. These guidelines should include the various types of graphic displays, as well as "best" type of display for a particular purpose.

b. Research and develop guidelines on the readability of graphic materials. Specifically, what elements in a particular type of graphic display make the display more understandable? Also, what are the functions of graphic displays and how can they best be utilized for teaching and/or learning?

c. Research and develop guidelines concerning design differences for materials which are to aid people in performing a task versus materials which are to aid people in learning a task.

Research to Provide Technical Writers with Information About Reader Characteristics and Purposes

A need exists to identify the characteristics and purposes of intended readers of different types of technical writing, and to get useful information of this sort to writers. Specific information about the important characteristics and purposes of readers of different types of material often does not get to writers before they begin their tasks; when any information is provided at all, it is likely to be too general to be of much help. Some writers may not even know what kind of information about readers could help them, and what kind they *would use* if it were available to them.

To provide concrete guidance to writers, we need to know *if* and *when* rewriting can make a reliable difference in reader comprehension and performance. Put another way, a model is needed that takes into account the several categories of factors that can influence the effectiveness of rewritten material. Characteristics of written material and of readers are clearly important; characteristics of the situation under which the material is used and/or tested may have more subtle but equally important effects. One such way is by influencing reader motivation and giving or not giving it a chance to affect reader behavior (e.g., through provision of liberal reading and/or testing time). One prediction from this model that has been successfully tested, is that improved readability will produce relatively greater information gain from a less preferred than from a more preferred content. A number of implications are suggested by such a finding. For examples, priority might well be given to: rewriting low-interest (versus high-interest) contents; putting the better writers to work on low-interest contents; rewriting material for low-motivation jobs and for low-motivation readers; most careful rewriting of the first parts of courses; et cetera. Further work on such a model might well be of value for rewriting tasks in the Armed Services, since the staggering volume of technical material makes essential a priority approach to the problem.

Research on Methods for Assisting the Writer, As Writer and As Editor

Discussion suggested that technical manual production, and particularly rewriting, are low priority activities. The growing power, and the attendant willingness, to reject unsatisfactory technical material should improve matters, since contractors must give increasing attention to technical manual production under these circumstances. But even then, few who have done much rewriting would suggest that it is as rewarding as writing an original draft. The computer seems to offer potential help, especially if the functions assigned it are those the writer does not like to do, or do well, and that are compatible with computer capabilities. For example, computers can be made to do language recording and storage tasks easily and efficiently, but language decision tasks only with difficulty and poorly, as the machine translation literature clearly shows.

Humans, on the other hand, have just the opposite proclivities. Consequently, computer aid to the writer should focus on the former, and should have the added characteristics of mapping onto printing and/or publishing functions already being done by computer or slated for computer handling later. For example, more and more typesetting is being done by computer. Readability formula programs could easily be tied into the output, as could cloze-formatting programs. In addition, records could be kept of the frequency of occurrence of technical terms (or other content terms) for use later in decisions on rewriting. In time, programs that provide alternative wordings (sets of synonyms, phrases, or even definitions) could be displayed for the writer at points likely to cause difficulty. But the *writer* would decide among them, *not* the computer — this the writer can do relatively easily and well, but the computer cannot. Similarly, a computer program could be developed relatively readily to flag difficult sentence constructions as long as it was not asked to transform such difficult constructions. One of the suggestions to grow out of the Conference was that separating the writing and editing functions performed by a writer could make the development of computer assistance even more logical. For example, the computer could help the writer in his or her *writing* role by storing and providing information useful in decisions concerning the organizing and sequencing of content. On the other hand, the computer could help the writer in his or her *editing* (rewriting) role by storing alternative wordings or flagging points of possible difficulty, i.e., "keeping track" (as noted above) of possible points of difficulty and reminding the writer of their existence as he goes over materials.

Understanding and Improving Information Systems Within the Armed Services

A persistent theme in Conference discussions was the practical purposes for which various written documents can and/or should be used, and the nature of the skills demanded of the readers in these uses. This appears to be partly an empirical question and partly a problem for systems analysis. In order to understand the current information systems within the Armed Services, a vigorous program is needed to determine the information demands in the various tasks that have to be carried out by a variety of occupational specialties, both on the job and in training to learn the job. It is also necessary to get reliable data as to how these information needs are met at present, what skills these information sources require from users, and the attitudes of users as well as training and command personnel toward the importance of requisite literacy and communications skills.

A prudent approach would be to supplement this important empirical work with a systematic analysis of what information systems ought to be like in order to be both efficient and effective. These analyses ought to consider the tasks that must be performed, as well as production, update, and distribution problems, and the impact of microfilm on technical

material production and use. It would probably be worthwhile to model and to simulate systems for various trade-offs between training and the use of performance aids. The problem of archival versus documentation for use deserves a closer examination for various practical circumstances. Manuals and accompanying materials specifically designed to teach were less emphasized at the Conference than manuals designed to facilitate job performance once an individual is on the job. The implication of this may be that the manuals are difficult to use in a training setting and force the instructor to improvise or that trainees are required to learn from manuals not especially designed for teaching them. Whether or not a differentiation between job manuals on the one hand and teaching/training manuals on the other can make a significant improvement in training should be a matter for study.

The rapid and accelerating growth in the amount of technical publications which are deemed necessary for hardware systems of increasing technical sophistication suggest the need for a major re-examination of documentation practices. Location of needed information, accessibility and updating, and physical bulk are only part of the problem. What appears to be also needed is better understanding of verbal and conceptual preparation necessary to operate effectively within a given information system. There are also indications that further developments in making text more readable and better suited to human purposes would not be fully utilized unless the institutions and organizations which now produce technical documents are modified. Concrete proposals for such organizational (and procedural) changes did not emerge from the Conference, though there seemed to be some agreement that the promulgation of simple rules for judging the acceptability of documents was not sufficient even if these are incorporated in regulations and procurement standards.

Finally, an information system that is totally integrated from beginning to end could provide feedback to the people who design the hardware and software systems, so that they might consider the documentation requirements (and resulting communications skills required of readers) at the same time that they make decisions about equipment structure, function, and maintenance.

REFERENCES AND BIBLIOGRAPHY

Academic Remedial Training San Diego: Comprehensive document. San Diego, California: Recruit Training Command, Academic Remedial Training Unit, 1975.

Aquino, M. R. The validity of the Miller-Coleman readability scale. *Reading Research Quarterly*, Spring 1969, 342-357.

Army Materiel Command, *Equipment publications readability program.* Paper presented at the AMC/NSIA Equipment Manuals Symposium, June 1970.

Atkinson, R. C. Ingredients for a theory of instruction. *American Psychologist*, 1972, 27, 921-931.

Ayoub, M. A., Cole, J. L., Sakala, M. K, & Smillie, R. J. *Job performance aids: Assessment of needs.* North Carolina State University, Raleigh, North Carolina, October 1974.

Ball, J. & Jamison, D. Computer-assisted instruction for dispersed populations: System cost models. *Instructional Science*, 1973, 1, 469-502.

Battelle Memorial Institute. *Interim report on the Office of Economic Opportunity experiment in educational performance contracting.* New York; Author, January 1972.

Battelle, R. B., Brown, H. D., Kruzie, P. G., Marshall, T. H., Moll, K. D., Paskert, P. F., & Radovic, M. *Analysis of some potential manpower policies for the all volunteer Navy.* Menlo Park, California: Stanford Research Institute, June 1973.

Berkshire, J. R. *Field evaluation of a troubleshooting aid.* AFPTRC-TR-54-24. Air Force Personnel and Training Research Center, Lackland AFB, Texas: June 1954.

Bialek, H. M., Taylor, J. E., & Hauke, R. N. *Instructional strategies for training men of high and low aptitude.* HumRRO Technical Report 73-10. Alexandria, Virginia: Human Resources Research Organization, April 1973.

Biersner, R. J. *Reading grade levels of Navy rate training manuals and non-resident career courses.* Chief of Naval Education and Training Support Report 2-75, May 1975.

Bloom, B. S. Time and learning. *American Psychologist*, 1974, 29, 682-688.

Bloomfield, L. & Barnhart, C. L. *Let's read: A linguistic approach.* Detroit, Michigan: Wayne State University Press, 1961.

Bormuth, J. R. Readability: A new approach. *Reading Research Quarterly*, 1966, 1, 79-131.

Bormuth, J. R. *Development of readability analysis*. Final Report, Project No. 7-0052, Contract No. OEC-3-7-070052-0326. U. S. Office of Education, Bureau of Research, 1969.

Bormuth, J. R. Reading literacy: Its definition and assessment. In J. B. Carroll and J. S. Chall (Eds.), *Toward a literate society*. New York: McGraw-Hill, 1975.

Braddock, W. Army targets its audience. In *Proceedings of Symposium on Technical Publications and the New User Profile, (8-9 October 1975)*, National Security Industrial Association Publication, No. 166, 1975.

Braid, T., Jr. Comparative analysis of three validation programs. In *Proceedings of Symposium on Technical Publications and the New User Profile, (8-9 October 1975)*, National Security Industrial Association Publication, No. 166, 1975.

Brinton, W. C. *Graphic methods of presenting facts*. New York: Engineering Magazine, 1961.

Bunde, G. R. *An effectiveness evaluation between manual and automated readability counting techniques*. Chief of Naval Education and Training Support Report 5-75, August 1975.

Burr, H. L. *Education in the early Navy*. Doctoral dissertation, Temple University. Philadelphia, Pennsylvania: Published by the author, 1939.

Burt, C. *A psychological study of typography*. Cambridge University Press, 1959.

Carroll, J. B. Development of native language skills beyond the early years. In C. E. Reed (Ed.), *The learning of languages*. New York: Appleton-Century-Crofts, 1971.

Carver, R. P. *Measuring the reading ability levels of Navy personnel*. Technical Report under contract to the Office of Naval Research, Psychological Sciences Division. Washington, D. C.: American Institutes for Research, October 1973. (a)

Carver, R. P. *Measuring the reading difficulty levels of Navy training manuals*. Technical Report under contract to the Office of Naval Research, Psychological Sciences Division. Washington, D. C.: American Institutes for Research, October 1973. (b)

Carver, R. P. *New techniques for measuring and improving reading comprehension.* American Institutes for Research, Technical Report No. 1, Contract No. N00014-72-C-0240 for Office of Naval Research, Personnel and Training Research Programs, February 1973. (c)

Carver, R. P. *The relationship between intellectual capacity and reading competency.* A Research Proposal. Silver Spring, Maryland: American Institutes for Research, September 1973. (d)

Carver, R. P. *Improving reading comprehension: Measuring readability.* American Institutes for Research, Final Report, Contract No. N00014-72-C-0240 for Office of Naval Research, Personnel and Training Research Programs, May 1974.

Caylor, J. S., Sticht, T. G., Fox, L., & Ford, J. *Methodologies for determining requirements of military occupational specialties.* HumRRO Technical Report 73-5. Alexandria, Virginia: Human Resources Research Organization, March 1973.

Chenzoff, A. P., Mallory, W. J., & Joyce, R. P. *Guidance and specification for the preparation of fully proceduralized job aids for organizational and intermediate maintenance of electronic sub-systems.* AFHRL-TR-71-23. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, June 1971. AD 731 144.

Cofer, C. N. *Adjustment to recruit training: A study of the effects of recruit preparatory training.* Technical Bulletin 54-22. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, December 1954.

Coke, E. U. *Reading rate, readability, and variations in task-induced processing.* Murray Hill: Bell Laboratories, (undated).

Coleman, M. & Liau, T. A computer readability formula designed for machine scoring. *Journal of Applied Psychology,* 1975, 60, 283-284.

Continental Army Command, Headquarters. Letter, dated 29 January 1973, Subject: Reading Level of Publications.

Couch, P. E. *Educational emphasis in civilian conservation camps of the Seventh Corps area.* Unpublished doctoral dissertation, Indiana University, 1944.

Cummings, D. E. Enlisted training in the Navy. *United States Naval Proceedings,* 1929, 55, 878-886.

Curran, T. E. *Review of technical manual readability and comprehensibility.* Navy Personnel Research and Development Center Technical Report, in preparation.

Curriculum for recruit preparatory training. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, April 1953.

Dailey, J. T. *The effect of screening and allocation policies on the quality of entering airmen.* HRRC-RB-51-19. Lackland AFB, Texas: Air Training Command, Human Resources Research Center, August 1951.

Dale, E. & Chall, J. S. A formula for predicting readability and instructions. *Educational Research Bulletin*, 1948, 27, 11-20, 28, 37-54.

Dansereau, D. F., Actkinson, T. R., Long, G. L., & McDonald, B. A. *Learning strategies: A review and synthesis of the current literature.* AFHRL-TR-74-70. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, December 1974. AD A007 722.

Dansereau, D. F., Evans, S. H., Actkinson, T. A., Long, C. L. *Factors relating to the development of optimal instructional information sequences.* AFHRL-TR-73-51(II). Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, July 1974. AD 783 843.

Dansereau, D. F., Evans, S. H., Wright, A. D., Long, G. L., & Actkinson, T. R. *Factors related to developing instructional information sequences: Phase I.* AFHRL-TR-73-51 (I). Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, March 1974. AD 777 832.

Dansereau, D. F., Long, G. L., McDonald, B. A., & Actkinson, T. R. *Learning strategy inventory development and assessment.* AFHRL-TR-75-40. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, June 1975.

Dansereau, D. F., Long, G. L., McDonald, B. A., Actkinson, T. R., Ellis, A. M., Collins, K., Williams, S., & Evans, S. H. *Effective learning strategy training program: Development and assessment.* AFHRL-TR-75-41. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, June 1975.

Davis, R. M. *Effective technical communications, mechanical description-experiment II.* TR-65-1, Wright-Patterson AFB, Ohio: Air Force Institute of Technology, School of Engineering, 1965. AD 631 238.

Davis, R. M. *Effective technical communications mechanical description-experiment III.* TR-66-15. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, School of Engineering, 1966. AD 641 251.

Davis, R. M. *Effective technical communications copy preparation and reproduction. Experiment II.* 4-74-AFIT-EN. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, February 1974. (a) AD 776 786.

Davis, R. M. *Effective technical communications, copy preparation and reproduction - motivation.* 6-74-AFIT-EN, TR-74-6. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, July 1974. (b) AD 783 842.

Davis, R. M. *Effective technical communications: Expression-copy preparation-motivation.* 7-74-AFIT-EN Technical Report. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, School of Engineering, August 1974. (c) AD 784 226.

Davis, R. M. Does expression make a difference - three experiments. *Proceedings of the 22nd International Technical Communications Society.* Washington, D. C.: Society for Technical Communications, 1975.

DeGuelle, J. R. *USAFI Achievement Test III diagnosis of reading deficiencies for Air Force personnel.* Unpublished manuscript, Aerospace Defense Command, January 1975.

Department of the Air Force. *Guide for Air Force writing.* AFM 11-3. Washington, D. C., April 1956 (Now out of print).

Department of the Army. *Marginal man and military service, a review.* January 1965.

Department of the Army. *Survey estimate of subject area of most reading for self improvement and primary reason for reading: Army male personnel.* OPOPM Report No. 29-68-E. Office of Personnel Operations, Personnel Management Development Office, February 1968.

Drury, C. M. *The history of the Chaplain Corps, United States Navy.* Washington, D. C.: Government Printing Office, 1949.

Duffy, T. M., Carter, J. C., Fletcher, J. D., & Aiken, E. G. *Language skills: A research and development plan for the naval service.* Special Report 76-3. San Diego, California: Navy Personnel Research and Development Center, October 1975.

Duffy, T. M. & Nugent, W. *Reading ability and attrition.* Unpublished manuscript, Navy Personnel Research and Development Center, 1975.

Elliott, T. K. *The effect of format and detail of job performance aids on performing simulated troubleshooting tasks.* AMRL-TR-65-154. Wright-Patterson AFB, Ohio: Aerospace Medical Research Laboratory, November 1965. AD 629 992.

Elliott, T. K. *Development of fully proceduralized troubleshooting routines.* AMRL-TR-67-152. Wright-Patterson AFB, Ohio: Aerospace Medical Research Laboratory, November 1967. AD 664 076.

Elliott, T. K. & Joyce, R. P. *An experimental comparison of procedural and conventional electronic troubleshooting.* AFHRL-TR-68-1. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, November 1968. AD 681 510.

Emeson, D. L. & Wulff, J. J. *The relationship between "What is learned" and "How it's taught".* ML-TM-57-32. Lowry AFB, Colorado: Air Force Personnel and Training Research Center, December 1957.

England, A. O. Getting your message across by plain talk. *Journal of Applied Psychology*, 1950, 34, 182-185.

England, A. O. Influence of 'plain talk' on AMC communications. *Journal of Applied Psychology*, 1951, 35, 381-382.

Evans, H. *Editing and design, a five-volume manual of English, typography and layout.* Heinemann, 1972-1975.

Fattu, N. A., Fay, L. C., D'Amico, L. A., & Standlee, L. S. *Observations on Navy literacy training.* Memorandum Report. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, March 1954.

Fattu, N. A., Mech, E., & Standlee, L. S. *A review of literacy training programs in the armed services during World War II.* Technical Bulletin 53-4. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, December 1953.

Fattu, N. A. & Standlee, L. S. *Analysis of reading difficulty of selected Navy materials.* Technical Bulletin 54-3. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, March 1954.

Fink, C. *Technical manuals for maintenance support: A maintenance rationale, some research findings, and some projections.* HumRRO Professional Paper 37-67. Alexandria, Virginia: Human Resources Research Organization, August 1967.

Fisher, A. H., Jr. *Army "new standards" personnel: Effect of remedial literacy training on performance in military service.* HumRRO Technical Report 71-7. AFHRL TR 71-13. Alexandria, Virginia: Air Force Human Resources Laboratory, Manpower Development Division, April 1971.

Flanagan, J. C. The critical incident technique. *Psychological Bulletin*, 1954, 51, 327-358.

Flesch, R. F. A new readability yardstick. *Journal of Applied Psychology*, 1948, 32, 221-233.

Flesch, R. F. *How to test readability.* New York: Harper, 1951.

Flew, A. G. N. *Thinking about thinking.* Fontana, 1975.

Flyer, E. S. *Factors relating to discharge for unsuitability among 1956 airman accessions to the Air Force.* WACC-TN-59-201. Lackland AFB, Texas, December 1959. AD 230 758.

Foley, J. P., Jr. *Task analysis for job performance aids and related training.* AFHRL-TR-72-73. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, October 1972.

Foley, J. P., Jr. *Job performance aids for UH-1H helicopter: Typcial fully proceduralized aids for non-troubleshooting tasks.* AFHRL-TR-75-28 (II). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, June 1975. (a)

Foley, J. P., Jr. *Job performance aids for UH-1H helicopter: Typical fully proceduralized aids for troubleshooting tasks.* AFHRL-TR-75-28 (III). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, June 1975. (b)

Folley, J. D., Jr. *Research problems in the design of performance aids.* ASD-TR-61-548. Wright-Patterson AFB, Ohio: Aeronautical Systems Division, October 1961. (a) AD 270 866.

Folley, J. D., Jr. *A preliminary procedure for systematically designing performance aids.* ASD-TR-61-550. Wright-Patterson AFB, Ohio: Aeronautical Systems Division, October 1961. (b) AD 270 868.

Folley, J. D., Jr., & Altman, J. W. *Guide to design of electronic equipment for maintainability.* WADC-TR-56-218. Dayton, Ohio: Wright Air Development Center, 1956.

Folley, J. D., Jr., & Elliott, T. K. *A field survey of electronic maintenance technical data.* AMRL-TR-67-159. Wright-Patterson AFB, Ohio: Aerospace Medical Research Laboratory, November 1967. AD 666 990.

Folley, J. D., Jr., Joyce, R. P., Mallory, W. J., & Thomas, D. L. *Fully proceduralized job performance aids: Volume I--Draft specification for organizational maintenance.* AFHRL-TR-71-53 (I). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, AFSC, December 1971. (a)

Folley, J. D., Jr., Joyce, R. P., Mallory, W. J., & Thomas, D. L. *Fully proceduralized job performance aids, Volume II--Developer's Handbook.* AFHRL-TR-71-53 (II). Brooks AFB, Texas: Air Force Human Resources Laboratory, AFSC, December 1971. (b)

Folley, J. D., Jr., & Munger, S. J. *A review of the literature on design of informational job performance aids.* ASD-TR-61-549. Wright-Patterson AFB, Ohio: Aeronautical Systems Division, October 1961. AD 270 867.

Folley, J. D., Jr., & Shettel, H. H. *Tryout of a preliminary procedure for systematically designing performance aids.* AMRL-TR-62-20. Wright-Patterson AFB, Ohio: Aerospace Research Laboratory, April 1962. AD 283 605.

Fox, W. L., Taylor, J. E., & Caylor, J. S. *Aptitude level and the acquisition of skills and knowledges in a variety of military training tasks.* HumRRO Technical Report 69-6. Alexandria, Virginia: Human Resources Research Organization, May 1969.

Gates, A. I. & MacGinitie, W. H. *Gates-MacGinitie Reading Tests, Survey D.* New York: Teachers College Press, 1965.

Gebhard, R. M. *Development of a training program and job aids for maintenance of electronic communication equipment.* HumRRO Technical Report 70-19. Alexandria, Virginia: Human Resources Research Organization, December 1970.

Ginzberg, E. & Bray, D. W. *The uneducated.* New York: Columbia University Press, 1953.

Goffard, J. S. *An experimental evaluation of a basic education program in the Army.* Technical Report 28 April 1956. Washington, D. C.: The Human Resources Research Office, 1956.

Goldberg, S. *Army training of illiterates in World War II.* Contributions to Education, Teachers College, Columbia University No. 966. New York: Bureau of Publications, Teachers College, Columbia University, 1951.

Gordon, M. A. & Bottenberg, R. A. *Prediction of unfavorable discharge by separate educational levels.* PRL-TDR-62-5. Lackland AFB, Texas: Air Force Personnel Research Laboratory, Aerospace Medical Division, April 1962. AD 284 802.

Gordon, M. A. & Flyer, E. S. *Predicted success of low-aptitude airmen.* PRL-TDR-62-14. Lackland AFB, Texas: Air Force Personnel Research Laboratory, Aerospace Medical Division, August 1962. AD 290 545.

Gragg, D. B., Kieselbach, D. J., Murphy, W. F., Peckham, R. E., & Heller, H. *The 14-week exploratory study of marginal-airman basic training: Comparison of proficiency of 8-week and 14-week training groups.* AFPTRC-TN-55-10. Lackland AFB, Texas: Air Force Personnel and Training Research Center, June 1955.

Gray, W. S. & Leary, B. E. *What makes a book readable.* Chicago: University of Chicago Press, 1935.

Guilford, J. P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

Guilford, J. P. & Hoepfner, R. *The analysis of intelligence.* New York: McGraw-Hill, 1971.

Gunderson, E. K. & Ballard, K. B. *Prediction of delinquency in naval recruits.* Prepared under ONR Contract NONR 1535 (00). San Diego, California: United States Retraining Command, undated.

Gunn, R. L. *An empirical study of the job components check list.* AFPTRC-TN-56-123. Lackland AFB, Texas: Air Force Personnel and Training Research Center, October 1956.

Gunning, R. *The technique of clear writing.* (Rev. ed.) New York: McGraw-Hill, 1968.

Hagen, E. P. & Thorndike, R. L. *A study of the World War II Navy careers of illiterates sent through literacy training.* Research Report, Contract NONR-644 (00). Washington, D. C.: Bureau of Naval Personnel, Classification and Survey Research Branch, April 1953.

Hamblin, C. L. *Fallacies.* Methuen, 1970.

Harman, D. Illiteracy: An overview. *Harvard Education Review*, 1970, 40, 226-243.

Harrod, F. S. *Enlisted men in the United States Navy, 1899-1939.* Unpublished doctoral dissertation, Northwestern University, June 1973.

Heath, S. R., Jr. The teaching of reading in a military setting. *The Training Bulletin*, 1956, 43, 150-156.

Hegg, J. O. & Weaver, D. O. Evaluating the reading difficulty of training literature. *ATC Training Analysis and Development Informational Bulletin*, Vance AFB, Oklahoma, 1952, 3, 19-21.

Highlander, R. W. Technical publication: Problems for the Army user. *In Proceedings of Symposium on Technical Publications and the New User Profile, (8-9 October 1975)*, National Security Industrial Association Publication No. 166, 1975.

Hodges, C. I. *Special training programs.* Unpublished Report. Bureau of Naval Personnel, November 1964.

Hoehn, A. J. *The development of training programs for first enlistment personnel in electronics maintenance MOSs. Part III. How to design the handbook materials.* George Washington University Publication, 1960. AD 631 069.

Hoehn, A. L. & Aukes, L. E. *Experiments on a guide for troubleshooting a simple electrical system* Unpublished Staff Memorandum, Air Force Personnel and Training Research Center, Maintenance Laboratory, Lowry AFB; Colorado, 3 February 1958.

AD-A034 730   HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA   F/G 5/10
READING AND READABILITY RESEARCH IN THE ARMED SERVICES.(U)
SEP 76   T G STICHT, D W ZAPF   N00014-76-C-0312

UNCLASSIFIED   HUMRRO-FR-WD(CA)-76-4   NL

4 OF 4
AD-A
034 730

END
DATE
FILMED
3-7-77
NTIS

Hoehn, A. J. & Lumsdaine, A. A. *Design and use of job aids for communicating technical information.* AFPTRC-TR-58-7. Lackland AFB, Texas: Air Force Personnel and Training Research Center, January 1958. AD 152 109.

Hoehn, A. J. & Saltz, E. *Determination of the behavioral content of a troubleshooting guide.* TARL Lab. Note 56-19 (unpublished). Chanute AFB, Illinois: Air Force Personnel and Training Research Center, Training Aids Research Laboratory, March 1956.

Hoehn, A. J. & Wardell, W. C. *Development of an experimental troubleshooting guide for the F-86D electronic fuel control system.* AFPTRC-ML-TM-57-25. Lowry AFB, Colorado: Air Force Personnel and Training Research Center, Maintenance Laboratory, December 1957.

Hoiberg, A., Hysham, C. J., & Berry, N. H. *The neuropsychiatric implications of illiteracy: 20 years later.* Report No. 74-20. San Diego, California: Neuropsychiatric Research Unit, February 1974.

Horton, R. E. Basic training for 'slow learning' airmen. *Training Analysis and Development Informational Bulletin,* c1951, 29-32.

Huff, K. H. & Smith, E. A. *Reliability, baseline data, and instructions for the automated readability index.* AFHRL-TR-70-14. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, October 1970.

Jablonski, W. A. *Improving the readability of maintenance manuals.* ABS-TP-1. Applied Behavioral Sciences, Inc., 1971.

Jensen, A. *Genetics & education.* New York: Harper and Row, 1972.

Johnson, K. H., Relova, R. P., Jr., & Stafford, J. P. *An analysis of the relationship between readability of Air Force procedural manuals and discrepancies involving non-compliance with the procedures.* Master's thesis. SLSR-25-72B. Wright-Patterson AFB, Ohio: Air Force Institute of Technology, School of Systems and Logistics, September 1972. AD 750 917.

Joyce, R. P. & Chenzoff, A. P. *Improving job performance aids through condensation, dual-level presentation, promotion of learning and entry by malfunction symptoms.* AFHRL-TR-74-12. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, March 1974.

Joyce, R. P., Chenzoff, A. P., Mulligan, J. F., & Mallory, W. J. *Fully proceduralized job performance aids: Draft military specification for organizational and intermediate maintenance.* AFHRL-TR-73-43 (I). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, December 1973. (a)

Joyce, R. P., Chenzoff, A. P., Mulligan, J. F., & Mallory, W. J. *Fully proceduralized job performance aids: Handbook for JPA developers.* AFHRL-TR-73-43 (II). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, December 1973. (b)

Joyce, R. P., Chenzoff, A. P., Mulligan, J. F., & Mallory, W. J. *Fully proceduralized job performance aids: Handbook for JPA managers and training specialists.* AFHRL-TR-73-43 (III). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, December 1973. (c)

Joyce, R. P., Folley, J. D., Jr., & Elliott, T. K. *Fully proceduralized job performance aids, Volume III--JPA Manager's Handbook.* AFHRL-TR-71-53 (III). Brooks AFB, Texas: Air Force Human Resources Laboratory AFSC, December 1971.

Judy, C. J. *Appraisal of education requirements for airman specialties.* WADD-TN-60-264. Lackland AFB, Texas: Personnel Laboratory, Wright Air Development Division, December 1960. AD 252 264.

Kent, W. P. *A longitudinal evaluation of the adult basic education program.* Falls Church, Virginia: Systems Development Corporation, November 1973.

Kent, W. P., Bishop, R. L., Byrnes, M. L., Frankel, S. M., Herzog, J. K., & Griffith, W. S. *Job-related adult basic education.* Volume II. Falls Church, Virginia: System Development Corporation, February 1971.

Kern, R. P. & Sticht, T. G. *Guidebook for the development of Army training literature: Rationale and policy implications from interviews with Army writers.* Unpublished Interim Report, U.S. Army Research Institute, March 1974

Kern, R. P., Sticht, T. G., Welty, D., & Hauke, R. N. *Guidebook for the development of Army training literature.* Arlington, Virginia: U.S. Army Research Institute, 1976, in press.

Kincaid, J. P., Fishburne, R. P., Jr., Rogers, R. L., & Chissom, B. S. *Derivation of new readability formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy enlisted personnel.* Research Branch Report 8-75. Naval Technical Training Command, February 1975.

Kincaid, J. P., Yasutake, J. Y., & Geiselhart, R. *Use of the automated readability index to assess comprehensibility of Air Force technical orders.* SEG-TR-67-47. Wright-Patterson AFB, Ohio: Systems Engineering Group, Aeronautical Systems Division, AFSC, November 1967.

Klare, G. R. *The measurement of readability.* Ames, Iowa: Iowa State University Press, 1963.

Klare, G. R. Assessing readability. *Reading Research Quarterly*, 1974-1975, 10, (1).

Klare, G. R. *A manual for readable writing.* Glen Burnie, Maryland: REM Company, 1975.

Klare, G. R., Gustafson, L. M., Mabry, J. E., & Shuford, E. H. The relationship of immediate retention of technical training material to career preferences and aptitudes. *Journal of Educational Psychology*, 1955, 46, 321-329.

Klare, G. R., Mabry, J. E., & Gustafson, L. M. *The relationship of verbal communication variables to immediate and delayed retention and to acceptability of technical training materials.* AFPTRC-TR-54-103. Lackland AFB, Texas: Air Force Personnel and Training Research Center, December 1954.

Klare, G. R., Mabry, J. E., & Gustafson, L. M. The relationship of patterning (underlining) to immediate retention and to acceptability of technical material. *Journal of Applied Psychology*, 1955(a), 39, 40-42.

Klare, G. R., Mabry, J. E., & Gustafson, L. M. The relationship of human interest to immediate retention and to acceptability of technical material. *Journal of Applied Psychology*, 1955(b), 39, 92-95.

Klare, G. R., Mabry, J. E., & Gustafson, L. M. The relationship of style difficulty to immediate retention and to acceptability of technical material. *Journal of Educational Psychology*, 1955(c), 46, 287-295.

Klare, G. R., Nichols, W. H., & Shuford, E. H. The relationship of typographic arrangement to the learning of technical training material. *Journal of Applied Psychology*, 1957, 41(1), 41-45.

Klare, G. R., Rowe, P. P., St. John, M. G., & Stolurow, L. M. *A first version of a computer-aided revising, editing, and translating system (CARET I).* Technical Report No. B.2.1. Harvard Computer-Aided Instruction Laboratory, 1969.

Klare, G. R., Shuford, E. H., & Nichols, W. H. The relationship of style difficulty, practice, and ability to efficiency of reading and to retention. *Journal of Applied Psychology*, 1957, 41(4), 222-226.

Klare, G. R., Shuford, E. H., & Nichols, W. H. The relation of format organization to learning. *Educational Research Bulletin, Ohio State University*, 1958, 37, 39-45.

Klare, G. R., Sinaiko, H. W., & Stclurow, L. M. Cloze procedure: A convenient readability test for training materials and translations. *International Review of Applied Psychology*, 1972, 21, 77-106.

Klein, D. *The Army writer: A guide to military writing.* Harrisburg, Pennsylvania: The Military Service Publishing Company, 1946.

Klesch, J. J. K. Improving troubleshooting manuals. In *Proceedings of Symposium on Technical Publications and the New User Profile, (8-9 October 1975)*, National Security Industrial Association Publication No. 166, 1975.

Lambert, J. V. & Siegel, A. I. Psycholinguistic determinants of readability. In A. I. Siegel & J. R. Burkett (Eds.), *Application of the structure-of-intellect and psycholinguistic concepts to reading comprehensibility measurement.* AFHRL-TR-74-49. Lowry AFB, Colorado: Air Force Human Resources Laboratory, 1974.

Langley, H. D. *Social reform in the United States Navy, 1798-1862.* Urbana, Illinois: University of Illinois Press, 1967.

Lecznar, W. B. *Some aptitude data on Air Force enlisted accessions.* PRL-TDR-62-10. Lackland AFB, Texas: Personnel Research Laboratory, Aerospace Medical Division, June 1962. AD 289 874.

Lewis, B. N., Horabin, I. S., & Gane, C. P. *Flow-charts, logical trees and algorithms for rules and regulations.* CAS Occasional paper 2. London: Her Magesty's Stationary Office, 1967.

Lorge, I. Predicting readability. *Teacher's College Record*, 1944, 45, 404-409.

Losee, J. E., Allen, R. H., Stroud, J. W., & Ver Hulst, J. *A study of the Air Force Maintenance Technical Data System.* AMRL-TDR-62-85. Behavioral Sciences Laboratory, Aerospace Medical Division. Air Force Systems Command, August 1962.

Luce, S. B. The manning of a Navy and mercantile marine. *United States Naval Institute Proceedings*, 1874, 1, 17-37.

Luce, S. B. Naval Training. *United States Naval Institute Proceedings*, 1890, 16, 367-430

Luckiesh, M. & Moss, F. K. *Reading as a visual task.* New York: Van Nostrand, 1942.

Lumsdaine, A. A. & Hoehn, A. J. *Preliminary study of audio job-performance aids.* Unpublished Staff Memorandum, Air Force Personnel and Training Research Center, Maintenance Laboratory, Lowry AFB, Colorado, 15 January 1958.

MacCaslin, E. F., Woodruff, A. B., & Baker, R. A. *An improved advanced individual training program for armor,* HumRRO Technical Report 59. Alexandria, Virginia: Human Resources Research Organization, December 1959.

Macdonald-Ross, M. *The quality of course materials in evaluation of teaching in higher education.* University of London Teaching Methods Center, 1976. (a)

Macdonald-Ross, M. *The quality of Open University course materials.* A contribution to the Ford Foundation conference on curriculum design, IET. The Open University, Milton Keynes, England: The Open University, 1976. (b)

Macdonald-Ross, M. *How numbers are shown: A review of the graphic presentation of quantitative data.* Preprints from author at IET. Milton Keynes, England: The Open University, 1976. (c)

Macdonald-Ross, M. & Waller, R. *Open University texts: criticism and alternatives.* Milton Keynes, England: Institute of Educational Technology, The Open University, 1975. (a)

Macdonald-Ross, M. & Waller, R. Criticism, Alternatives and Tests: A conceptual framework for improving typography. *Programmed Learning and Educational Technology,* 1975 (b), 12(2), 75-83.

Macdonald-Ross, M. & Waller, R. The Transformer. *Penrose Annual,* in press.

Madden, H. L. & Tupes, E. C. *Estimating reading ability level from the AEQ general aptitude index.* PRL-TR-66-1. Lackland AFB, Texas: Personnel Research Laboratory, Aerospace Medical Division, February 1966. AD 632 182.

Mastropaolo, S., Carp, A., & Erdmann, R. L. *A study of the relative effects of six-week and twelve-week experimental basic training programs on a sample of limited-aptitude airmen: Part III Eight-month follow-up comparison.* AFPTRC-TR-54-37. Lackland AFB, Texas: Air Force Personnel and Training Research Center, September 1954.

Mastropaolo, S., Carp, A., Erdmann, R. L., & Schmid, J. Jr. *A study of the relative effects of six-week and twelve-week experimental basic training programs on a sample of limited-aptitude airmen: Part I. Basic training analyses. Part II. Six-week follow-up analyses.* AFPTRC-TR-54-36. Lackland AFB, Texas: Air Force Personnel and Training Research Center, September 1954.

McGoff, R. M. & Harding, F. D. *A report on literacy training programs in the armed services.* AFHRL-TR-73-69. Air Force Human Resources Laboratory, Manpower Development Branch, 1973.

McReynolds, J. *Aptitude levels in the enlisted manpower pool of the Air Force.* WADC-TN-58-63(I) Part I, Appendix, WADC-TN-58-63(II) Part II. Lackland AFB, Texas: Wright Air Development Center, September 1958. AD 151 048.

McReynolds, J. *Use of the Airman Qualifying Examination to predict completion of basic training.* PRL-TDR-64-9. Lackland AFB, Texas: Air Force Personnel Research Laboratory, April 1964. AD 442 047.

Miller, R. B. *A suggested guide to the preparation of handbooks and job instructions.* ML-TM-56-15. Lowry AFB, Colorado: Air Force Personnel and Training Research Center, Maintenance Laboratory, May 1956.

Milton, H. H. Let's improve our writing, not measure it. *ATC Training Analysis and Development Informational Bulletin*, 1953, 4, 13-17.

Mockovak, W. P. *An analysis of Air Force reading improvement programs: Results of USAF Survey Number 73-89.* AFHRL-TR-73-54. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, January 1974. (a) AD 775 047.

Mockovak, W. P. *An investigation and demonstration of methodologies for determining the reading skills and requirements of Air Force career ladders.* AFHRL-TR-73-53. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, January 1974. (b)

Mockovak, W. P. *Literacy skills and requirements in Air Force career ladders.* AFHRL-TR-74-90. Lowry AFB, Colorado: Air Force Human Resources Laboratory, December 1974. (c)

Mullen, P. A., & Joyce, R. P. *Demonstration of fully proceduralized job performance aids and matching training.* AFHRL-TR-74-69. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, August 1974.

Nath, C. J. *Preliminary analysis of AF reading program.* Briefing presented at the Air Force Reading Literacy Conference, Washington, D.C., May 1975.

Naval Air Systems Command. *Technical writers guide* (Preliminary) NAOO-25-700. NAVAIRSYSCOM, 1976.

Nelson, M. J. *The Nelson Reading Test.* (Rev. ed.) Boston, Massachusetts: Houghton Mifflin Co., 1962.

Nelson, M. J., & Denny, E. C. *The Nelson-Denny Reading Test.* (Rev. by J. I. Brown.) Boston, Massachusetts: Houghton Mifflin Co., 1960.

Neurath, M. Isotype. *Instructional Science*, 1975, 3, 127-50.

Neurath, O. *International picture language.* Kegan Paul, 1936.

Newman, S. E. *Method for conducting a survey of practices in the use of maintenance handbooks.* Unpublished Staff Memorandum, Air Force Personnel and Training Research Center, Maintenance Laboratory, Lowry AFB, Colorado, November 1957 (a)

Newman, S. E. *On identifying effective handbook techniques.* AFPTRC-ML-TM-57-26. Lowry AFB, Colorado: Air Force Personnel and Training Research Center, Maintenance Laboratory, December 1957. (b)

Niblack, A. P. The enlistment, training, and organization of crews for our new ships. *United States Naval Institute Proceedings*, 1891, 17, 3-49.

Northcutt, N., Selz, N., Shelton, E., Nyer, L., Hickok, D., & Humble, M. *Adult functional competency: A summary.* Austin, Texas: University of Texas, March 1975.

Olson, D. R. A review of: Toward a literate society. *Proceedings of the National Academy of Education*, 1975, 2, 109-178.

Paterson, D. G. & Tinker, M. A. *How to make type readable.* New York: Harper, 1940.

Peerson, V. L. Evaluating the readability of training literature. *ATC Training Analysis and Development Informational Bulletin*, 4, Keesler AFB, Mississippi, 1953, 48-53.

Phillips, J. E. *The use of training time information in planning and evaluating manpower programs.* Report prepared by the California Occupational Analysis Field Center, April 1970.

Pieper, W. J., Catrow, E. J., Swezey, R. W., & Smith, E. A. *Automated apprenticeship training (AAT): A systematized audio-visual approach to self-paced job training.* AFHRL-TR-72-20. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, April 1973.

Plag, J. A. *Predicting the effectiveness of naval enlistees. (Tables and Summary of a Report.)* San Diego, California: Naval Medical Neuropsychiatric Research Unit, May 1968.

Plag, J. A., & Hardacore, L. E. *The validity of age, education, and GCT score as predictors of two-year attrition among naval enlistees.* Report Number 64-15. San Diego, California: Navy Medical Neuropsychiatric Research Unit, June 1964.

Post, T. J. *Guide for selecting formats and media for maintenance information*, Draft Report. Fall Church, Virginia: BioTechnology, Inc., October 1975.

Post, T. J. & Price, H. E. *Requirements and criteria for improving reading comprehension of technical manuals.* Contract N00024-74-C-5426, for Naval Sea Systems Command, Technical Publications Branch. Fall Church, Virginia: BioTechnology, Inc., November 1974.

Potter, F. H. A school for bluejackets. *Outlook*, 1913, 114, 694-705.

Potter, F. H. A repair shop for men. *Outlook*, 1918, 120, 539-540.

Pritsker, A. A. B. *The Q-GERT user's manual*. West Lafayette, Indiana. Pritsker & Associates, Inc., May 1974.

Pritsker, A. A. B. & Happ, W. W. GERT: Part I - Fundamentals. *Journal of Industrial Engineering*, 1966, XVII (5), 267.

Pritsker, A. A. B. & Sigal, C. E. *The GERT III Z user's manual*. West Lafayette, Indiana: Pritsker & Associates, Inc., August 1974.

Project One Hundred Thousand: Characteristics and performance of "new standards" men. Washington, D. C.: Department of Defense, Manpower and Reserve Affairs, March 1969.

Pyke, R. L. *Report on the legibility of print*. London: Her Majesty's Stationary Office, 1926.

Raven, J. C. *Standard progressive matrices: Sets A, B, C, D, & E*. London, England: H. K. Lewis & Co., Ltd., 1958.

Rees, D. W. *Guide to design of Air Force check-list publications*. WADC-TR-59-758. Wright-Patterson AFB, Ohio: Air Research and Development Command, December 1959.

Rees, D. W., & Copeland, N. K. *The effects of serial position in check-list design*. WADC-TR-59-552. Wright-Patterson AFB, Ohio: Air Research and Development Command, September 1959.

Rees, D. W. & Kama, W. N. *Size of tabs: A factor in handling of guides and check-lists*. WADC-TR-59-158. Wright-Patterson AFB, Ohio: Air Research and Development Command, March 1959. AD 213 595.

Research for Better Schools. *A pilot model remedial education program*. Philadelphia, Pennsylvania: Author, 1974.

Rogers, J. P. & Harris, J. S. *Preparation of MAINTRAIN trouble-shooting manuals*, Working Paper. Fort Bliss, Texas: U. S. Army Air Defense Human Research Unit, October 1964.

Rogers, J. P., & Thorne, H. W. *The development and evaluation of an improved electronics troubleshooting manual*. HumRRO Technical Report 65-1, MAINTRAIN V. Alexandria, Virginia: Human Resources Research Organization, March 1965. AD 614 606.

Ronan, W. W., & Prien, E. P. *Perspectives on the measurement of human performance*. New York: Appleton-Century-Crofts, 1971.

Ross, C. S. Literacy training in the Navy. *School and Society*, 1946, 63, 203-204.

Ross, D. A. *Comprehensibility evaluation of technical manuals.* WADC TN-59-442. Wright-Patterson AFB, Ohio: Wright Air Development Center, July 1959. AD 118 235.

Rubenstein, H. & Aborn, M. Learning, prediction, and readability. *Journal of Applied Psychology*, 1958 42(1), 28-32.

Sakala, M. K. *Effects of format structure, sex and task familiarity on the comprehension of procedural instructions.* Unpublished thesis, Department of Psychology, North Carolina State University, October 9, 1975.

Sellers, L. *The simple subs book, a manual for sub-editors (and would-be sub-editors) on newspapers and house journals.* Pergamon, 1968.

Sellman, W. S. *Effectiveness of experimental training materials for low ability airmen.* AFHRL-TR-70-16. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, June 1970.

Sellman, W. S. Reducing the literacy demands of training materials as one means of increasing learning in low-ability personnel. *The Journal of Experimental Education*, 1972, 41(2), 54-59.

Shanley, D. C. & Smith, R. G., Jr. *An evaluation of the language arts aspect of the basic training program.* AFPTRC-TN-55-36. Lackland AFB, Texas: Air Research and Development Command, November 1955.

Sharon, A. *What do adults read?* Research Bulletin 72-57. Princeton, New Jersey: Educational Testing Service, December 1972.

Shennum, W. A., Aiken, E. G., & Thomas, G. S. *Uses of time-compressed speech in a reading remediation program: Some exploratory tests.* Technical Report TR 76-13. San Diego, California: Navy Personnel Research and Development Center, September 1975.

Showel, M. & Brennan, M. *A survey of user attitudes towards Army training literature.* ARI Research Problem Review. U. S. Army Research Institute, 1976, in press.

Shriver, E. L. *Fully proceduralized job performance aids: Guidance for performing behavioral analyses of tasks.* AFHRL-TR-75-38. Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, June 1975.

Shriver, E. L. & Foley, J. P., Jr. *Job performance aids for UH-1H helicopter: Controlled field tryout and evaluation.* AFHRL-TR-75-28(I). Wright-Patterson AFB, Ohio: Air Force Human Resources Laboratory, Advanced Systems Division, June 1975.

Shriver, E. L. & Trexler, R. C. *A description and analytic discussion of ten new concepts for electronics maintenance.* HumRRO Technical Report 66-23. Alexandria, Virginia: Human Resources Research Organization, December 1966.

Siegel, A. I. & Bergman, B. A. Readability/comprehensibility as related to the structure-of-intellect model. In A. I. Siegel & J. R. Burkett (Eds.), *Application of the structure-of-intellect and psycholinguistic concepts to reading comprehensibility measurement.* AFHRL-TR-74-49. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, 1974.

Siegel, A. I. & Burkett, J. R. (Eds.) *Application of structure-of-intellect and psycholinguistic concepts to reading comprehensibility measurement.* AFHRL-TR-74-49. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, September 1974.

Siegel, A. I., Federman, P. J., & Burkett, J. R. *Increasing and evaluating the readability of Air Force written materials.* AFHRL-TR-74-28. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, August 1974.

Siegel, A. I., Lambert, J. V., & Burkett, J. R. *Techniques for making written material more readable/comprehensible.* AFHRL-TR-74-47. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, August 1974.

Siegel, A. I., Lautman, M. R., & Burkett, J. R. Reading grade level adjustment and auditory supplementation as techniques for increasing textual comprehensibility. *Journal of Educational Psychology,* 1974, 66, 895-902.

Siegel, A. I., & Wolf, J. J. Computer analysis of textual comprehensibility. In O. Lecarme & R. Lewis (Eds.), *Computers in Education.* Amsterdam: North Holland Publishing Company, 1975.

Singer, H. IQ is and is not related to reading. Paper presented at the annual meeting of the International Reading Association, Denver, Colorado, May 1973.

Smillie, R. J., Edsall, J. C., Sakala, M. K., & Ayoub, M. A. *Assessment and evaluation of job performance aids: A simulation approach.* Draft, North Carolina State University, 1975.

Smith, A. D. W. *Generic skills for occupational training.* Prince Albert, Saskatchewan: Training Research and Development Station, 1973.

Smith, E. A. & Kincaid, J. P. Derivation and validation of the automated readability index for use with technical materials. *Human Factors*, 1970, 12(5), 457-464. AD 716 314. (Also identified as AFHRL-TR-70-47.

Smith, E. A. & Senter, R. J. *Automated readability index.* AMRL-TR-66-220. Wright-Patterson AFB, Ohio: Aerospace Medical Division, November 1967. AD 667 273.

Smith, H. L. Improved technical documentation and training. In T. C. Rowan (Ed.), *Proceedings, Invitational conference on improved information aids for technicians.* Sponsored by The Society for Applied Learning Technology and The Logistics Management Institute, May 1975.

Spearman, C. Theory of a general factor. *British Journal of Psychology*, 1946, 36, 117-131.

Special training program. Unpublished report, Bureau of Naval Personnel, 1951.

Standlee, L. S. *A follow-up comparison of three groups of Navy enlisted men: Marginal-and-illiterate, marginal-but-literate, and typical recruits.* Technical Bulletin 54-20. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, July 1954.

Standlee, L. S., Fattu, N. A., & Auble, J. D. *Data on daily performance of marginal-and-illiterate, marginal-but-literate, and typical recruits from three naval training centers.* Memorandum Report. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, July 1954. (a)

Standlee, L. S., Fattu, N. A., & Auble, J. D. *Frequency index of words appearing in four Navy publications.* Technical Bulletin 54-2. Washington, D. C.: Bureau of Naval Personnel, Personnel Analysis Division, January 1954. (b)

Steineman, J., Hooprich, E. A., Archibald, A. G., & Van Matre, N. H. *Development of a "wordsmanship" training course for marginal personnel.* Research Report SRR 71-17. San Diego, California: Naval Personnel and Training Research Laboratory, February 1971.

Stephenson, H. W. Plain English. *U.S.A. Combat Forces Journal*, 1950, 1, 38-39.

Stewart, V. T. *1st Quarter Report of Phase II Pilot Program: Recruit Remedial Reading Program.* Prepared under Marine Corps Contract Number M00243-74-C-0020. San Diego, California: Community Colleges, Midway Adult School, 1974.

Stewart, V. T. *2nd Quarterly Report of Recruit Remedial Reading Program, Phase III.* Prepared under Marine Corps Contract Number M00243-75-C-0015. San Diego, California: Community Colleges, Midway Adult School, July 1975.

Sticht, T. G. Learning by listening. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* New York: John Wiley & Sons, 1972.

Sticht, T. G. The acquisition of literacy by children and adults. Paper presented at the Second Delaware Symposium on Curriculum, Instruction, and Learning: The Acquisition of Reading. University of Delaware, June 1975. (a)

Sticht, T. G. *A program of Army functional job reading training: Development, implementation, and delivery system.* HumRRO-FR-WD (CA)-75-7. Alexandria, Virginia: Human Resources Research Organization, June 1975. (b)

Sticht, T. G. (Ed.) *Reading for working: A functional literacy anthology.* Alexandria, Virginia: Human Resources Research Organization, 1975. (c)

Sticht, T. G. & Beck, L. J. *Development of an experimental literacy assessment battery.* HumRRO-FR-WD-CA-76-5. Alexandria, Virginia: Human Resources Research Organization, June 1976.

Sticht, T. G., Beck, L. J., Hauke, R. N., Kleiman, G. M. & James, J. H. *Auding and reading: A developmental model.* AFHRL-TR-74-36. Lowry AFB, Colorado: Technical Training Division, Air Force Human Resources Laboratory, July 1974.

Sticht, T. G., Caylor, J. S., Fox, L. C., Hauke, R. N., James, J. H., Snyder, S. S., & Kern, R. P., *HumRRO's literacy research for the U.S Army: Developing functional literacy training.* HumRRO Professional Paper 13-73. Alexandria, Virginia: Human Resources Research Organization, December 1973.

Sticht, T. G., Caylor, J. S., Kern, R. P., & Fox, L. C. *Determination of literacy skill requirements in four military occupational specialties.* HumRRO Technical Report 71-23. Alexandria, Virginia: Human Resources Research Organization, November 1971.

Sticht, T. G., Fox, L. C., Hauke, R. N., & Zapf, D. W. *Reading in the Navy.* HumRRO FR-WD-CA-76-14. Alexandria, Virginia: Human Resources Research Organization, May 1976.

Stolurow, L. M. & Newman, J. R. A factional analysis of objective features of printed language presumably related to reading difficulty. *Journal of Educational Research*, 1959, 52(7), 243-251.

Sulit, R. A. *Navy technical manual system.* Pre-proposal conference for RFP N00600-76-R-5050. Technical Memorandum DT NSRDC-TM-186-75-3. Bethesda: Naval Ship Research and Development Center, 1975.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 1953, 30, 415-433.

Taylor, W. L. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology*, 1957, 41, 19-26.

Thouless, R. H. *Straight and crooked thinking.* English University Press, 1953.

Vernon, M. D. Learning from graphical materials. *British Journal of Psychology*, 1946, 36, 145-58.

Vernon, M. D. The visual presentation of factual material. *British Journal of Education Psychology*, 1950, 20, 174-185.

Vernon, M. D. Learning and understanding. *Quarterly Journal of Experimental Psychology*, 1951, 3, 19-23.

Vernon, M. D. The use and value of graphical material in presenting quantitative data. *Occupational Psychology*, 1952 (a), 26, 22-34.

Vernon, M. D. The use and value of graphical materials with a written text. *Occupational Psychology*, 1952 (b), 26, 96-100.

Vernon, M. D. Presenting information in diagrams. *AV Communication Review*, 1953, 1, 147-58.

Vernon, M. D. *The psychology of perception.* 4th ed. Methuen, 1962.

Vineberg, R., Sticht, T. G., Taylor, E. N., & Caylor, J. S. *Effects of aptitude (AFQT), job experience, and literacy on job performance: Summary of HumRRO Work Units UTILITY and REALISTIC.* HumRRO Technical Report 71-1. Alexandria, Virginia: Human Resources Research Organization, February 1971.

Vineberg, R. & Taylor, E. *Performance in four Army jobs by men at different aptitude (AFQT) levels: 3. The relationship of AFQT and job experience to job performance.* HumRRO Technical Report 72-22. Alexandria, Virginia: Human Resources Research Organization, August 1972.

Vitola, B. M., Mullins, C. J., & Brokaw, L. D. *Quality of the all-volunteer Air Force--1973.* AFHRL-TR-74-35. Lackland AFB, Texas: Air Force Human Resources Laboratory, Personnel Research Division, April 1974.

Vitola, B. M., Valentine, L. D., Jr., & Tupes, E. C. *Aptitude and educational data for Air Force enlistees, 1962 through 1965.* PRL-TR-67-8. Lackland AFB, Texas: Personnel Research Laboratory, August 1967. AD 664 035.

Weeden, E. T. *Academic remedial training.* Unpublished report, Naval Training Center, January 1975.

Williams, A. R., Jr., Siegel, A. I., & Burkett, J. R. *Readability of textual material--A survey of the literature.* AFHRL-TR-74-29. Lowry AFB, Colorado: Air Force Human Resources Laboratory, Technical Training Division, July 1974.

Wing, C. W., Jr. & Wallach, M. A. *College admissions and the psychology of talent.* New York: Holt, Rinehart, & Winston, 1971.

Wulff, J. J. & Newman, S. E. *Illustrative guidance for job-task analysis--maintenance of C97-C landing light.* Unpublished Staff Memorandum, Lowry AFB, Maintenance Laboratory, Air Force Personnel and Training Research Center, October 1956.

Young, K. & Jamison, D. T. *The economic benefits of schooling and reading competence.* Report RB 75-19. Princeton, New Jersey: Educational Testing Service, May 1975.

Zaccaria, M. *Effectiveness of the Basic Military Training Reading Proficiency Training Program.* Unpublished Air Force report, 1971.

APPENDIX A

BIOGRAPHICAL NOTES

Biographical Notes on the Presentors

JAMES R. BURKETT

Ron Burkett is Deputy Chief of the Instructional Technology Branch at the Technical Training Division of the Air Force Human Resources Laboratory. From 1968 to 1969 he was a Personnel Measurement Psychologist at the Personnel Research Division of AFHRL engaged in the development and revision of Air Force specialty knowledge tests. Since joining the Technical Training Division of AFHRL in 1969 as a Research Psychologist, he has conducted and supervised research in the areas of instructional technology, readability, training evaluation, adult literacy skills, on-the-job training, and instructional systems and materials. Dr. Burkett has authored and co-authored over 20 technical reports, articles, and papers and was the 1974 winner of the AFHRL Donald B. Haines award for technical achievement. He received his PhD in Experimental Psychology from the University of Oklahoma.

JOHN S. CAYLOR

John Caylor is a Senior Scientist with the Human Resources Research Organization (HumRRO). He has more than 13 years experience in research and development in military settings, including work in the determination of job reading requirements and research leading to a job functional literacy training program. He has taught at the University of California, Riverside, and served as Evaluation Specialist for the Appalachian Adult Education Center. Dr. Caylor received his PhD from the University of Michigan.

THOMAS E. CURRAN

Thomas Curran, of the Navy Personnel Research and Development Center in San Diego, California, is Project Director for the Center's participation in the Navy Technical Manual System research program. Previously, he was Assistant Professor of Psychology at California State University, Chico, and has served as an officer in the U.S. Navy. His professional specialties include the psychology of human learning, human development, and the psychology of language. Dr. Curran received his PhD in Educational Psychology from the University of Minnesota.

Preceding page blank

THOMAS M. DUFFY

Tom Duffy is a research psychologist at the Navy Personnel Research and Development Center. Prior to joining the Center in 1973, he was an Assistant Professor in the Department of Psychology at the University of Manitoba. His current research involves the analysis of the reading skills of low literate adults and the development of programs for training functional literacy in the military context. This work developed out of a career interest in human memory and comprehension processes and verbal learning. Publications have included investigation of imagery and verbal mnemonics in learning and memory, as well as the requirements for reading skills in the Navy. He received his PhD in 1969 in experimental psychology from the University of Illinois.

J. DEXTER FLETCHER

Dexter Fletcher is a member of the Navy Personnel Research and Development Center. His principal interests have been in the design, development, and evaluation of computer-assisted instruction with particular emphasis on curriculum areas related to natural language processing. Currently, he is developing and evaluating techniques for the use of intelligent instructional systems for the Navy. Dr. Fletcher received an MS in computer science and a PhD in educational psychology, both from Stanford University.

STEVEN D. GROFF

Steven Groff is a Captain in the U.S. Air Force and is assigned to the Technical Training Division of the Air Force Human Resources Laboratory. Prior to this assignment, he served as Technical Programs Monitor for technical training and flying training research at the Air Force Human Resources Laboratory, Headquarters, Brooks AFB, Texas. He is currently engaged in research in the areas of literacy training and comprehensibility measurement. He received his MA in Experimental/Industrial Psychology from the University of Tulsa in 1971.

JACK H. HILLER

Jack Hiller is a Senior Project Director for the Army Research Institute at the Presidio of Monterey, California. He is currently developing a reading comprehension test, and planning research to develop a computer system which would provide writers with diagnostic feedback. Previously, Dr. Hiller worked for the Southwest Regional Lab for Educational R&D, where he planned research in reading comprehension, and designed instruction to promote development of comprehension processes. He has also been an Assistant Professor of Educational Psychology at Southern Illinois University, Carbondale, Illinois, where he conducted extensive research in learning from text and lecture, and analysis of natural language data by computer. He has published numerous articles on using the computer to analyze text and transcribe speech and on verbal behavior and verbal learning. He is a member of the American Psychological Association, American Educational Research Association, and is listed in *Who's Who*. Dr. Hiller earned his PhD in Educational Psychology from the University of Connecticut at Storrs.

ROBERT C. JOHNSON

Robert Johnson is a research psychologist with the Air Force Human Resources Laboratory. His work in job performance aids research includes development of handbooks detailing procedures for producing proceduralized technical data and a user acceptance and useability study of job guide manuals. Prior to joining AFHRL, Mr. Johnson served as an Air Force maintenance officer, and later as a Human Factors Senior Engineer for the Lockheed-Georgia Company. He currently serves as a maintenance officer for an Ohio Air National Guard fighter group. He holds a BS degree in Psychology from Northeast Missouri State University, and is currently enrolled in graduate work at the University of Dayton.

WILLIAM G. MULLER

William Muller is employed at the Naval Air Technical Services Facility as the Human Factor Specialist of the Naval Air Systems Command, Technical Documentation Policy and Programs Office. In this position he directs and manages technical documentation research, development and evaluation. Mr. Muller also presently holds an additional position of Human Factor Consultant to the Naval Technical Manual System Program and as Human Factors Advisor to the NAVAIR Weapons Training Division (AIR-413). Prior work experience has been at the Naval Air Development Center, Warminster, Pa., where his primary interests were in Technical Training and Human Factors Engineering for Future Naval Weapon Systems.

RICHARD P. KERN

Richard Kern is a Senior Research Psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences. Prior to this, he was a Senior Research Scientist with the Human Resources Research Organization. His major research interests have been in applied research on design of training and training media and in factors affecting quality of performance on the job. He is currently conducting research on design of performance-oriented nonresident instructional systems and their supporting materials and research to develop job analysis methodologies which will permit identification of the communication elements of job tasks and assessment of their related skill requirements. He is a member of the American Psychological Association, the American Educational Research Association, and the International Reading Association. Dr. Kern received his PhD from the State University of Iowa.

ARTHUR I. SIEGEL

Arthur Siegel is Director of Applied Psychological Services, Inc. Previously he was a Research Associate at Fordham University; taught psychology at New York University, at Queens College, and at Temple University; and was a Project Director at the Institute for Research in Human Relations. His studies have included the development of methods for quantifying the intellective and motor load on the system operator, work method development, manpower planning and personnel allocation, performance aid development, training system development, human reliability prediction, maintainability analysis, display and control design, human transfer function analysis, and system evaluation. Dr. Seigel's experience in linguistics and readability includes four Air Force programs, at least five Navy programs, and one Army program. Dr. Siegel has also developed a number of special techniques for analyzing and defining the comprehensibility of technical materials. He has published over 200 professional articles and technical reports, two books, and has presented 36 papers before the American and Eastern Psychological Associations. Dr. Siegel is a member of several professional associations. He received the doctoral degree from New York University.

Biographical Notes on Civilian Consultants

JEANNE S. CHALL

Jeanne Chall is Professor of Education and Director of the Reading Laboratory, Graduate School of Education, Harvard University. She has conducted research on readability, vocabulary, and the psychology and teaching of reading. She is a Fellow in the American Psychological Association and the American Association for the Advancement of Science, a past president of the National Conference on Research in English, and is a member of the Board of the National Society for the Study of Education. Among her publications are the Dale-Chall Formula for Predicting Readability (with Edgar Dale); Readability: An Appraisal of Research and Application; Learning to Read: The Great Debate; and Toward a Literate Society, with John B. Carroll. Dr. Chall received her PhD from the Ohio State University.

SAMUEL GIBBON

Samuel Gibbon has been with the Children's Television Workshop since 1968, when he joined that organization as one of the original Producers of Sesame Street. Previously, he was Associate Producer and Writer of Captain Kangaroo for more than six years. In 1970 Mr. Gibbon undertook the development of The Electric Company, a television program offering supplemental reading instruction for second-, third-, and fourth-graders experiencing reading difficulties. He served as Executive Producer of that program until 1973. Since 1972 he has held an appointment as Lecturer in Education and Research Associate in the Center for Research in Children's Television at the Harvard University Graduate School of Education. His research interests focus on the effects of the properties of the television medium and the "syntax" of televised messages on children's developing cognitive structures. Recent publications include NATIONAL CONFERENCE ON VISUAL INFORMATION PROCESSING, (National Institute of Education, 1975); and "Sesame Street, The Electric Company, and Reading" (in J. Carroll and J. Chall, Eds., TOWARD A LITERATE SOCIETY, McGraw-Hill, Inc., 1975). Mr. Gibbon received his AB in English Literature from Princeton University.

ROBERT GLASER

Robert Glaser is University Professor of Psychology & Education and Co-Director of the Learning Research & Development Center at the University of Pittsburgh. He has been President of the American Educational Research Association (1971-72) and of the Educational Psychology Division of the American Psychological Association. During 1969-70 Professor Glaser was a Fellow at the Center for Advanced Study in the Behavioral Sciences in Palo Alto, California. In 1970 he received the AERA/American Educational Publishers Institute Award for Outstanding Educational Research in the fields of Instructional Materials. And, in 1973 he was elected to the National Academy of Education. Professor Glaser has served as a consultant to many national and international groups, including the Educational Subpanel of the President's Science Advisory Committee. His recent research and publication is in the area of cognitive processes involved in intellectual performance, and the design of school environments that are adaptive to individual differences. Dr. Glaser received his PhD in 1949 in psychological measurement and experimental psychology from Indiana University.

JOHN T. GUTHRIE

John Guthrie has been the Director of Research for the International Reading Association since September, 1974. Previously, he served as Kennedy Scholar in Education at the Kennedy Institute of the Johns Hopkins University. Concept formation and verbal learning in children were his early research interests. In the past five years he has studied basic perceptual and language processes in reading, and in particular the cognitive deficiencies of poor readers. Aspects of Reading Acquisition, a volume recently edited by Guthrie, presents a multidisciplinary perspective of learning to read. It will be published by the Johns Hopkins University Press early in 1976. A second edited volume, Cognition, Curriculum, and Comprehension, which focuses on theory and practice in the development of reading comprehension, is in preparation. Dr. Guthrie received his PhD in educational psychology from the University of Illinois.

GEORGE R. KLARE

George Klare is Professor of Psychology at Ohio University. Previously he was Dean of the College of Arts & Sciences at Ohio University. Earlier affiliations were with the Training Research Laboratory of the Department of Psychology at the University of Illinois, and the Test Division of the Psychological Corporation. Professor Klare has been a visiting scholar at Harvard University, State University of New York at Stonybrook, and in Europe as recipient of a Baker Award. He has served as consultant/lecturer for many organizations, and is listed in Who's Who in America. Professor Klare's continuing research interests focus on readability and design of texts. His publications include The Measurement of Readability (Iowa State University Press), Elementary Statistics - Data Analysis for the Behavioral Sciences (McGraw-Hill), and Know Your Reader (Hermitage House). Dr. Klare received his PhD in psychology from the University of Minnesota.

MICHAEL MACDONALD-ROSS

Michael Macdonald-Ross has been associated with the Open University in England since its start in 1969. From 1969 to 1971 he was a partner in the educational consultancy, Instructional Systems Associates, which helped the University to set up its course development system. Since 1971 he has been Senior Lecturer in the University's Institute of Educational Technology. His research interests are directed toward curriculum design and the way print is used to communicate complex information. Currently, Mr. Macdonald-Ross is Director of the Institute's Textual Communication Research Group. This group conducts research into the ways students learn from texts, and advises the University on ways to improve its printed teaching materials. Recent publications include papers on behavioral objectives and the structure of knowledge, a bibliography for textual communication, an analysis of Open University texts, and a review of research into typography. In addition, Mr. Macdonald-Ross is a Policy Editor of the Journal of Instructional Science and a Council Member of the Institute of Scientific and Technical Communicators. He received his BSc in zoology from University College, London.

ERNST Z. ROTHKOPF

Ernst Rothkopf is Head of the Learning and Instruction Research Department at Bell Laboratories. He joined the Bell research organization in 1958 after five years as a civilian scientist in the Air Force's Personnel & Training Research Center. Dr. Rothkopf was one of the first researchers in programmed instruction. His main interests throughout his career have been in learning theory, training techniques, human verbal learning, and particularly, learning from written instructional material. He is the author of over sixty publications and the editor of a book, Verbal Learning and the Technology of Written Instruction. Dr. Rothkopf has been a visiting or adjunct professor at New York University, Rutgers, and the University of California at Berkeley. He is on several national advisory committees including the National Commission on Performance-Based Education, and is a Fellow of the American Psychological Association, as well as a member in numerous professional organizations. Dr. Rothkopf received his PhD in experimental psychology from the University of Connecticut.

RICHARD L. VENEZKY

Richard Venezky has been on the faculty at the University of Wisconsin since 1965. He has been a member of the NCTE Commission on Reading, the IRA Committee on Psycholinguistics, and the Advisory Board of the Central Midwest Regional Laboratory's Language Program. He is also director pro tempore for the IRA Oral Language Staff Development Program. His current research includes observational studies of classroom organization (sponsored by the Israeli Ministry of Education), analysis and review of teacher training techniques for oral language development (sponsored by the International Reading Association), and diagnosis of prereading and early reading skills (sponsored by NIE). His publications during the past year include monographs on testing and teaching in reading, and a prereading skills program developed at the Wisconsin Research and Development Center. Professor Venezky is currently Professor and Chairman of the Computer Sciences Department, and a principal investigator in the Wisconsin Research and Development Center for Cognitive Learning. He received his PhD in linguistics from Stanford University.

THOMAS G. STICHT (Conference Chairperson)

Tom Sticht is a Senior Staff Scientist with the Human Resources Research Organization/Western Division. Prior to this, he was a NASA Postdoctoral Fellow at the University of Pittsburgh, and Assistant Professor in the Department of Psychology at the University of Louisville. In the Spring of 1975 he was Visiting Associate Professor at the Graduate School of Education, Harvard University. For several years, Dr. Sticht's research and development interests have centered on the functional literacy problems encountered by adults. This work has emphasized the improvement of readers through reading training, and the improvement of materials through design for readability and useability. Most of this work has been for Air Force, Army, and Navy sponsors, and has been brought together in two recent publications - Reading for Working: A Functional Literacy Anthology (HumRRO), and Auding and Reading: A Developmental Model (HumRRO with others). Dr. Sticht has been a consultant to local school districts, education laboratories, and government agencies, and is currently the Chairman of the Basic Education and Reading Committee of the International Reading Association. He received his PhD in experimental psychology from the University of Arizona.

APPENDIX B

LIST OF CONFEREES:

Conference on Reading and Readability Research in the Armed Services

NAVY

1. Dr. Marshall J. Farr, Director
   Personnel and Training Research Programs
   Office of Naval Research (Code 458)
   Arlington, VA 22217

2. Dr. Edwin Aiken
   Navy Personnel Research and Development Center (Code 306)
   San Diego, CA 92152

3. CDR Larry Beguin
   Bureau of Naval Personnel
   PERS-2S
   Washington, DC 20320

4. Dr. Thomas Curran
   Navy Personnel Research and Development Center (Code 306)
   San Diego, CA 92152

5. Dr. Thomas Duffy
   Navy Personnel Research and Development Center (Code 306)
   San Diego, CA 92152

6. Dr. Dexter Fletcher
   Navy Personnel Research and Development Center (Code 306)
   San Diego, CA 92152

7. Dr. Eugene E. Gloye
   Office of Naval Research
   Pasadena Branch Office
   1030 East Green Street
   Pasadena, CA 91101

8. Dr. Norman J. Kerr
   Head, Research Section
   Training Methods Evaluation and Research Branch (Code 341)
   Chief of Naval Technical Training
   NAS Memphis (75)
   Millington, TN 38054

9. Mr. William Muller
   NAVAIR Technical Services Facility
   NATSF
   700 Robbins Avenue
   Philadelphia, PA 19111

AIR FORCE

10. Dr. James R. Burkett
Air Force Human Resources Laboratory
Technical Training Division
Lowry Air Force Base, CO 80230

11. Dr. Ronald P. Carver
University of Missouri at Kansas City
Kansas City, MO 64110

12. CAPT Charles Curran
ATC/XPTT
Randolph AFB, TX 78148

13. CAPT Steven G. Groff
Air Force Human Resources Laboratory
Technical Training Division
Lowry AFB, CO 80230

14. Dr. Raymond L. Lewiski
Extension Course Institute
Educational Systems Branch
Gunter AFB, AL 36114

15. Dr. Arthur I. Siegel
Applied Psychological Services
404 E. Lancaster Avenue
Wayne, PA 19087

ARMY

16. Dr. J. E. Uhlaner
US Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

17. Dr. Richard D. Bloom
U.S. Army Research Institute Field Unit
P. O. Box 5787
Presidio of Monterey, CA 93940

18. Mr. R. T. Diduk
HQ TRADOC (ATPR)
DCS Personnel
Education Branch
Fort Monroe, VA 23651

19. Mr. Dennis F. Fisher
Behavioral Research Directorate
US Army Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005

20. Mr. George Fry
Education Specialist
Army-Wide Training Literature
U.S. Army Signal School
Fort Gordon, GA 30905

21. MAJ William Highlander
HQ TRADOC (ATTNG-AS)
Collective Training Branch, DCST
Literature Production Section
Fort Monroe, VA 23651

22. Dr. Jack H. Hiller
US Army Research Institute Field Unit
P.O. Box 5787
Presidio of Monterey, CA 93940

23. Dr. Otto Kahn
US Army Research Institute Field Unit
P.O. Box 5787
Presidio of Monterey, CA 93940

24. Dr. Joseph H. Kanner
HQ TRADOC (ATTNG)
DCS Training
Educational Advisor's Office
Fort Monroe, VA 23651

25. Dr. Richard P. Kern
US Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

26. Mr. Herman Rupp
Army-Wide Training Support Division
P.O. Box 5300
US Army Air Defense School
Fort Bliss, TX 79916

27. Mr. Jack J. Sternberg
Chief, US Army Research Institute Field Unit
P.O. Box 5787
Presidio of Monterey, CA 93940

CIVILIANS

28. Dr. Francis A. Cartier
Chief, Development Division
Office of Research and Development
Building 276
Defense Language Institute
Presidio of Monterey, CA 93940

29. Dr. Jeanne S. Chall
Harvard University
Reading Laboratory
Graduate School of Education
Roy E. Larsen Hall, Appian Way
Cambridge, MA 02138

30. Dr. Esther Coke
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

31. Mr. Fred Finch
California Test Bureau/McGraw-Hill
Del Monte Research Park
Monterey, CA 93940

32. Mr. Robert Fishburne
Human Factors Section
Environment and Energy Systems Department
Calspan Corporation
Buffalo, NY

33. Dr. Lawrence Frase
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

34. Mr. Samuel Y. Gibbon
Executive Producer
Children's Television Workshop
One Lincoln Plaza
New York, NY 10023

35. Dr. Robert Glaser
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260

36. Philip K. Glossa
Project Director
Sierra Learning Laboratory
P.O. Box 497
Jamestown, CA 95327

37. Dr. Burl Gray
Behavioral Sciences Institute
72 Fern Canyon Road
Carmel Highlands, CA 93921

38. Dr. Donald Ross Green
Director of Research
California Test Bureau/McGraw-Hill
Del Monte Research Park
Monterey, CA 93940

39. Dr. John T. Guthrie
International Reading Association
800 Barksdale Road
Newark, DE 19711

40. Dr. George R. Klare
Department of Psychology
Ohio University
35 Park Place
Athens, OH 45701

41. Mr. Glenn Kleiman
Psychology Department
Stanford University
Stanford, CA 94305

42. Mr. Michael Macdonald-Ross
The Open University
Institute of Educational Technology: Textual Research Unit
Walton Hall, Milton Keynes
MK7 6AA, England

43. Mr. Monte Penney
National Institute of Education
1200 19th Street NW (Room 815)
Washington, CD 20208

44. Dr. Ernst Z. Rothkopf
Learning and Instructional Processes Research Group
Bell Laboratories
Murray Hill, NJ 07971

45. Dr. Glenn Roudabush
California Test Bureau/McGraw-Hill
Del Monte Research Park
Monterey, CA 93940

46. Mr. H. E. Schlichting
Senior Engineer, Technical Data
McDonnell Aircraft Company
St. Louis, MO 63166

47. Dr. Ed Smith
Psychology Department
Stanford University
Stanford, CA 94305

48. Dr. Richard L. Venezky
University of Wisconsin
Computer Sciences Department
1210 West Dayton Street
Madison, WI 53706

49. Phyllis Weaver
Learning Research and Development Center
University of Pittsburgh
Pittsburgh, PA 15260

Human Resources Research Organization (HumRRO)/Western Division
P.O. Box 5787 Presidio of Monterey, CA 93940

50. Dr. Thomas G. Sticht, Conference Chairman

51. Mrs. Diana Welty Zapf, Conference Coordinator

52. Dr. John C. Caylor

53. Mr. Lynn C. Fox

54. Mr. Robert N. Hauke

55. Mr. Kent H. Huff

56. Mr. Lawrence J. Beck